OM protein - protein search, using sw model

Run on: November 10, 2003, 01:51:30 ; Search time 80 Seconds
(without alignments)
710.301 Million cell updates/sec

Title: US-09-868-547-4
Perfect score: 1886
Sequence: 1 MASSLNNGRKASEIFQGQAL.....FQSYKISPFTGYLSLIEIYP 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1886	100.0	358	21	AAY96584	Glycine max isofla
2	1012	53.7	350	21	AAY96587	Glycine max isofla
3	1012	53.7	351	21	AAY96585	Glycine max isofla
4	1006	53.3	348	21	AAY96586	Glycine max partia
5	997	52.9	350	21	AAY96583	Glycine max partia
6	636.5	33.7	347	20	AAY24396	Coptis japonica no
7	590.5	31.3	375	20	AAY05661	Maize caffeic O-me
8	580.5	30.8	350	20	AAY27183	(S)-3'-hydroxy-N-m
9	567	30.1	370	20	AAY05662	Maize caffeic O-me
10	540.5	28.7	366	20	AAY05663	Maize caffeic O-me
11	527	27.9	368	21	AAG40137	Arabidopsis thalia
12	527	27.9	382	21	AAG40136	Arabidopsis thalia
13	527	27.9	382	23	ABB93248	Herbicidally activ
14	521	27.6	368	21	AAG28589	Arabidopsis thalia
15	521	27.6	382	21	AAG28588	Arabidopsis thalia
16	521	27.6	383	21	AAG28587	Arabidopsis thalia
17	485	25.7	325	21	AAG31303	Arabidopsis thalia
18	485	25.7	325	23	ABB93247	Herbicidally activ
19	435	23.1	354	20	AAY05660	Maize caffeic O-me
20	434	23.0	296	21	AAG40138	Arabidopsis thalia
21	426	22.6	363	20	AAY43041	Safener binding pr
22	392.5	20.8	253	21	AAG31304	Arabidopsis thalia
23	390.5	20.7	330	21	AAG25274	Arabidopsis thalia
24	390.5	20.7	330	21	AAG25598	Arabidopsis thalia
25	390.5	20.7	330	21	AAG47428	Arabidopsis thalia
26	390.5	20.7	340	21	AAG25597	Arabidopsis thalia
27	390.5	20.7	340	21	AAG47427	Arabidopsis thalia
28	390.5	20.7	363	21	AAG25596	Arabidopsis thalia
29	390.5	20.7	363	21	AAG47426	Arabidopsis thalia
30	390.5	20.7	363	23	ABB93811	Herbicidally activ
31	390.5	20.7	364	14	AAR34764	OMTI translation p
32	379.5	20.1	365	15	AAR63203	Alfalfa COMT. Med
33	379.5	20.1	365	22	AAE12021	Medicago sativa ca
34	376.5	20.0	316	21	AAG25275	Arabidopsis thalia
35	375.5	19.9	222	21	AAG31305	Arabidopsis thalia
36	371.5	19.7	364	14	AAR34762	pPLC4 translation
37	371.5	19.7	365	20	AAY01133	Aspen bispecific O
38	371.5	19.7	365	23	AAU80015	5-hydroxyconiferal
39	369.5	19.6	365	19	AAW84132	Exemplery caffeic
40	367.5	19.5	314	14	AAR34763	OMT1.A and OMT1.B
41	367	19.5	371	23	AAE16516	Rice caffeic acid
42	365	19.4	368	19	AAW84128	An (iso) eugenol m
43	364.5	19.3	309	21	AAG25276	Arabidopsis thalia
44	364.5	19.3	365	23	AAE16510	Soybean caffeic ac
45	362	19.2	368	24	AAE29785	Clarkia breweri IE

ALIGNMENTS

RESULT 1
AAY96584
ID AAY96584 standard; Protein; 358 AA.
XX
AC AAY96584;
XX
DT 26-SEP-2000 (first entry)
XX
DE Glycine max isoflavone O-methyltransferase.
XX
KW Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW Flavonoid; flower colour; growth; pollination; irradiation.
XX
OS Glycine max.
XX
PN WO200037656-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US30338.
XX
PR 21-DEC-1998; 98US-0113190.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;
XX
DR WPI; 2000-442680/38.
DR N-PSDB; AAA29322.
XX
PT New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT transgenic plants and for immunological screening of cDNA libraries

```
XX
PS   Claim 10; Page 32-33; 39pp; English.
XX
CC   AAY96583-87 are (partial) isoflavone O-methyltransferases isolated from
CC   various soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c
CC   and ssl1c were prepared from soybean embryo (19 days after flowering),
CC   root, 8-day old root and seed (25 days after flowering).
CC   Isoflavone O-methyltransferase catalyses the first step in degradation
CC   of daidzein. Suppression of this enzyme will yield higher concentrations
CC   of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC   co-pigments in flower colour, stimulate pollen tube growth, attract
CC   pollenators, act as feeding deterrents and protect against UV
CC   irradiation in fruits and seeds. The cDNA and proteins can be used to
CC   isolate homologues, for immunological screening and for positive
CC   selection methods.
XX
SQ   Sequence   358 AA;

  Query Match             100.0%;  Score 1886;  DB 21;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-189;
  Matches  358;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358


RESULT 2
AAY96587
ID   AAY96587 standard; Protein; 350 AA.
XX
AC   AAY96587;
XX
DT   26-SEP-2000  (first entry)
XX
DE   Glycine max isoflavone O-methyltransferase.
XX
KW   Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW   Flavonoid; flower colour; growth; pollination; irradiation.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 89
FT                   /note= "Encoded by GAAGATCATGAA"
XX
PN   WO200037656-A2.
XX
PD   29-JUN-2000.
XX
PF   20-DEC-1999;   99WO-US30338.
XX
PR   21-DEC-1998;   98US-0113190.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Famodu OO,  McGonigle B,  Odell JT,  Fader GM,  Falco SC;
XX
DR   WPI; 2000-442680/38.
DR   N-PSDB; AAA29325.
XX
PT   New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT   transgenic plants and for immunological screening of cDNA libraries
XX
PS   Claim 10; Page 37-39; 39pp; English.
XX
CC   AAY96583-87 are (partial) isoflavone O-methyltransferases isolated from
CC   various soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c
CC   and ssl1c were prepared from soybean embryo (19 days after flowering),
CC   root, 8-day old root and seed (25 days after flowering).
CC   Isoflavone O-methyltransferase catalyses the first step in degradation
CC   of daidzein. Suppression of this enzyme will yield higher concentrations
CC   of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC   co-pigments in flower colour, stimulate pollen tube growth, attract
CC   pollenators, act as feeding deterrents and protect against UV
CC   irradiation in fruits and seeds. The cDNA and proteins can be used to
CC   isolate homologues, for immunological screening and for positive
CC   selection methods.
XX
SQ   Sequence   350 AA;

  Query Match             53.7%;  Score 1012;  DB 21;  Length 350;
  Best Local Similarity   54.9%;  Pred. No. 2.4e-97;
  Matches  196;  Conservative   69;  Mismatches   84;  Indels    8;  Gaps    5;

Qy          3 SSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVS 62
              :|:|| :|  |:|:||:|||  | | :   |||| |:| ||||| :|:  :||: |:|||
Db          1 ASMNN-QKEIELFEGQSLLYMQLYGHLRPMCLKWAVQLGIPDIIQNHA--KPISLSDLVS 57

Qy         63 ILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIE-AYALTAASELLVKSSELSLAPMVE 121
              ||:||     ||  ||:|||||  ||   ||:: |  |||| ||:||| ||:  |:|||
Db         58 TLQIPPANAAFVQRFMRFLAHNGIFE---IHESQELTYALTPASKLLVNSSDHCLSPMVL 114

Qy        122 YFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLN 181
              | :|     :: |  |:  || :||| :  ||  |  : |:| |   |||||| ||:::
Db        115 AFTDPLRNVKYHHLGEWIRGEDPSVFETAHGTSAWGLLEKNPEYFSLFNEAMASDSRIVD 174

Qy        182 LAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFV 241
              || ::|  |||||:|:|||||||| ||:|||:|||||||:||: |:|||||:|:|||:||
Db        175 LALKNCTSVFEGLDSMVDVGGGTGTTARIICDAFPKLKCVVLDLPHVVENLTGTNNLSFV 234

Qy        242 GGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD 301
              |||||  ||:|||||| ||||| | :|:|||: |:::|| :  :|||::|| ||||  |
Db        235 GGDMFNSIPQADAVLLKWVLHNWTDENCIKILQKCRDSISSKGNSGKVIIIDAVINEKLD 294

Qy        302 ERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :  :|: || :|: |   :||:|| |::||:||:|||| ||| |  |: ||||:||
Db        295 DPDMTQTKLSLDIIM-LTMNGRERTEKEWKQLFIEAGFKHYKIFPIFGFRSLIEVYP 350


RESULT 3
AAY96585
ID   AAY96585 standard; Protein; 351 AA.
XX
AC   AAY96585;
XX
DT   26-SEP-2000  (first entry)
XX
DE   Glycine max isoflavone O-methyltransferase.
XX
KW   Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW   Flavonoid; flower colour; growth; pollination; irradiation.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
```

```
FT   Misc-difference 90
FT                   /note= "Encoded by GAAGATCATGAA"
XX
PN   WO200037656-A2.
XX
PD   29-JUN-2000.
XX
PF   20-DEC-1999;   99WO-US30338.
XX
PR   21-DEC-1998;   98US-0113190.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Famodu OO,  McGonigle B,  Odell JT,  Fader GM,  Falco SC;
XX
DR   WPI; 2000-442680/38.
DR   N-PSDB; AAA29323.
XX
PT   New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT   transgenic plants and for immunological screening of cDNA libraries
XX
PS   Claim 10; Page 34-35; 39pp; English.
XX
CC   AAY96583-87 are (partial) isoflavone O-methyltransferases isolated from
CC   various soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c
CC   and ssl1c were prepared from soybean embryo (19 days after flowering),
CC   root, 8-day old root and seed (25 days after flowering).
CC   Isoflavone O-methyltransferase catalyses the first step in degradation
CC   of daidzein. Suppression of this enzyme will yield higher concentrations
CC   of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC   co-pigments in flower colour, stimulate pollen tube growth, attract
CC   pollenators, act as feeding deterrents and protect against UV
CC   irradiation in fruits and seeds. The cDNA and proteins can be used to
CC   isolate homologues, for immunological screening and for positive
CC   selection methods.
XX
SQ   Sequence   351 AA;

  Query Match             53.7%;  Score 1012;  DB 21;  Length 351;
  Best Local Similarity   54.9%;  Pred. No. 2.4e-97;
  Matches  196;  Conservative   69;  Mismatches   84;  Indels    8;  Gaps    5;

Qy          3 SSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVS 62
              :|:|| :|  |:|:||:|||  | | :   |||| |:| ||||| :|:  :||: |:|||
Db          2 ASMNN-QKEIELFEGQSLLYMQLYGHLRPMCLKWAVQLGIPDIIQNHA--KPISLSDLVS 58

Qy         63 ILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIE-AYALTAASELLVKSSELSLAPMVE 121
               ||:||     ||  ||:|||| ||   ||:: |  |||| ||:||| ||:  |:|||
Db         59 TLQIPPANAAFVQRFMRFLAHNGIFE---IHESQELTYALTPASKLLVNSSDHCLSPMVL 115

Qy        122 YFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLN 181
              | :|     :: |  |:  || :||| : ||  |  : |:| |    |||||| ||::::
Db        116 AFTDPLRNVKYHHLGEWIRGEDPSVFETAHGTSAWGLLEKNPEYFSLFNEAMASDSRIVD 175

Qy        182 LAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFV 241
              || ::|  |||||:|:||||||||| ||:|||:|||||||:||: |:|||||:|:|||:||
Db        176 LALKNCTSVFEGLDSMVDVGGGTGTTARIICDAFPKLKCVVLDLPHVVENLTGTNNLSFV 235

Qy        242 GGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD 301
              |||||  ||:|||||||| ||||| | :|:|||: |:::|| :  :|||::|| ||||  |
Db        236 GGDMFNSIPQADAVLLKWVLHNWTDENCIKILQKCRDSISSKGNSGKVIIIDAVINEKLD 295

Qy        302 ERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :  :|: || :|: |   :||:|| |::||:|:|:||||: ||| |  |: ||||:||
Db        296 DPDMTQTKLSLDIIM-LTMNGRERTEKEWKQLFIEAGFKHYKIFPIFGFRSLIEVYP 351


RESULT 4
AAY96586
ID   AAY96586 standard; Protein; 348 AA.
XX
AC   AAY96586;
XX
DT   26-SEP-2000  (first entry)
XX
DE   Glycine max partial isoflavone O-methyltransferase.
XX
KW   Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW   Flavonoid; flower colour; growth; pollination; irradiation.
XX
OS   Glycine max.
XX
PN   WO200037656-A2.
XX
PD   29-JUN-2000.
XX
PF   20-DEC-1999;   99WO-US30338.
XX
PR   21-DEC-1998;   98US-0113190.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Famodu OO,  McGonigle B,  Odell JT,  Fader GM,  Falco SC;
XX
DR   WPI; 2000-442680/38.
DR   N-PSDB; AAA29324.
XX
PT   New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT   transgenic plants and for immunological screening of cDNA libraries
XX
PS   Claim 10; Page 35-36; 39pp; English.
XX
CC   AAY96583-87 are (partial) isoflavone O-methyltransferases isolated from
CC   various soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c
CC   and ssl1c were prepared from soybean embryo (19 days after flowering),
CC   root, 8-day old root and seed (25 days after flowering).
CC   Isoflavone O-methyltransferase catalyses the first step in degradation
CC   of daidzein. Suppression of this enzyme will yield higher concentrations
CC   of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC   co-pigments in flower colour, stimulate pollen tube growth, attract
CC   pollenators, act as feeding deterrents and protect against UV
CC   irradiation in fruits and seeds. The cDNA and proteins can be used to
CC   isolate homologues, for immunological screening and for positive
CC   selection methods.
XX
SQ   Sequence   348 AA;

  Query Match             53.3%;  Score 1006;  DB 21;  Length 348;
  Best Local Similarity   54.7%;  Pred. No. 1e-96;
  Matches  192;  Conservative   64;  Mismatches   91;  Indels    4;  Gaps    3;

Qy          9 RKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPP 68
              :|| |:|:||:|||  | | :   |||| |:| ||||| :|:  :||| |:||| ||:||
Db          1 QKAMELFEGQSLLYMQLYGHLRPMCLKWAVQLGIPDIIQNHA--KPITLSDLVSTLQIPP 58

Qy         69 TKTRQVQSLMRYLAHNGFFEIVRIHDNIE-AYALTAASELLVKSSELSLAPMVEYFLEPN 127
              :|   ||  ||:|||:| |:|    |: | ||||| ||:||| |:  |:|||    :|
Db         59 SKAGFVQQFMRFLAHDGIFDIRESQDDHELAYALTPASKLLVSCSDHCLSPMVRMNTDPL 118

Qy        128 CQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDC 187
                  ::    |:  || || | : || ||  : |:|     |||||| ||:|::|| ::|
Db        119 LMTTYHHFGEWIRGEDPTVHETAFGTSFWGLLEKNPTQMSLFNEAMASDSRMVDLALKNC 178

Qy        188 NWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFK 247
              |||||:|:||||||| ||:|||||||||:||: |:|||||:|:|||:|||||||
Db        179 TSVFEGLDSMVDVGGGTGTTARIICEAFPKLKCVVLDLPHVVENLTGTNNLSFVGGDMFN 238

Qy        248 CIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTE 307
              |: |||||| |||||| :|:|||: ||::|| :   |||::|| :|||  |:  :|
Db        239 SFPQTDAVLLKWVLHNWNDENCIKILKKCKDSISSKGNRGKVIIIDIIINEKLDDPDMTR 298

Qy        308 LKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              || :|: |:  :||:|| |::||::|:|||||  || |  |: ||||:||
Db        299 TKLSLDIVMS-TMNGRERSEKEWKQMFIEAGFQHCKIFPIFGFRSLIELYP 348
```

RESULT 5
AAY96583
ID AAY96583 standard; Protein; 350 AA.
XX
AC AAY96583;
XX
DT 26-SEP-2000 (first entry)
XX
DE Glycine max partial isoflavone O-methyltransferase.
XX
KW Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW Flavonoid; flower colour; growth; pollination; irradiation.
XX
OS Glycine max.
XX
PN WO200037656-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US30338.
XX
PR 21-DEC-1998; 98US-0113190.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;
XX
DR WPI; 2000-442680/38.
DR N-PSDB; AAA29321.
XX
PT New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT transgenic plants and for immunological screening of cDNA libraries
XX
PS Claim 10; Page 30-31; 39pp; English.
XX
CC AAY96583-87 are (partial) isoflavone O-methyltransferases isolated from
CC various soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c
CC and ssl1c were prepared from soybean embryo (19 days after flowering),
CC root, 8-day old root and seed (25 days after flowering).
CC Isoflavone O-methyltransferase catalyses the first step in degradation
CC of daidzein. Suppression of this enzyme will yield higher concentrations
CC of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC co-pigments in flower colour, stimulate pollen tube growth, attract
CC pollenators, act as feeding deterrents and protect against UV
CC irradiation in fruits and seeds. The cDNA and proteins can be used to
CC isolate homologues, for immunological screening and for positive
CC selection methods.
XX
SQ Sequence 350 AA;

Query Match 52.9%; Score 997; DB 21; Length 350;
Best Local Similarity 54.1%; Pred. No. 9e-96;
Matches 191; Conservative 65; Mismatches 93; Indels 4; Gaps 3;

```
Qy          7 NGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQV 66
              : :|| |:|:||:|||  | | :   |||| |:| ||||| :|:  :||| |:||| ||:
Db          1 DNQKAMELFEGQSLLYMQLYGHLRPMCLKWAVQLGIPDIIQNHA--KPITVSDLVSTLQI 58

Qy         67 PPTKTRQVQSLMRYLAHNGFFEIVRIHDNIE-AYALTAASELLVKSSELSLAPMVEYFLE 125
               |:|   ||  ||:|||:| |:|    |: | ||||| ||:|||  |:  |:|||    :
Db         59 SPSKAGFVQQFMRFLAHDGIFDIRESQDDHELAYALTPASKLLVSCSDHCLSPMVRMNTD 118

Qy        126 PNCQGAWNQLKRWVHEEDLTVPEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFR 185
              |     ::    |:  || || | : || ||  : |:|     |||||| ||:|::|| : :
Db        119 PLLMTTYHHFGEWIRGEDPTVHETAFGTSFWGLLEKNPTQMSLFNEAMASDSRMVDLALK 178

Qy        186 DCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDM 245
              :|  |||||:|:|||||||| ||||||||||||||:|| |:|||||:|:|||:||||||
Db        179 NCTSVFEGLDSMVDVGGGTGTTAKIICEAFPKLKCVVLDLPHVVENLTGTNNLSFVGGDM 238

Qy        246 FKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQV 305
              |   |: |||||| ||||||| :|:|||: ||::|| :   |||::|| :|||  |:  :
Db        239 FNSFPQTDAVLLKWVLHNWNDENCIKILKKCKDSISSKGNKGKVIIIDIIINEKLDDPDM 298

Qy        306 TELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              |  || :|: |:  :||:|| |::||::|:||||:  || |  |: ||||:||
Db        299 TRTKLSLDIVMS-TMNGRERSEKEWKQMFIEAGFKHCKIFPIFGFRSLIELYP 350
```

RESULT 6
AAY24396
ID AAY24396 standard; Protein; 347 AA.
XX
AC AAY24396;
XX
DT 20-SEP-1999 (first entry)
XX
DE Coptis japonica norcoclaurine 6-0-methyltransferase.
XX
KW Coptis japonica; norcoclaurine 6-0-methyltransferase; 6-OMT;
KW secondary metabolite; coclaurine; reticuline.
XX
OS Coptis japonica.
XX
PN JP11178577-A.
XX
PD 06-JUL-1999.
XX
PF 24-DEC-1997; 97JP-0355045.
XX
PR 24-DEC-1997; 97JP-0355045.
XX
PA (MITC) MITSUI PETROCHEM IND CO LTD.
XX
DR WPI; 1999-437310/37.
DR N-PSDB; AAX90136.
XX
PT New norcoclaurine 6-O-methyl-transferase - useful for preparation of
PT secondary plant metabolite
XX
PS Claim 1; Page 11-12; 27pp; Japanese.
XX
CC The present sequence represents norcoclaurine 6-O-methyltransferase
CC (6-OMT). Also described is a method for the preparation of the secondary
CC metabolite of a plant derived from coclaurine or reticuline.
XX
SQ Sequence 347 AA;

Query Match 33.7%; Score 636.5; DB 20; Length 347;
Best Local Similarity 38.9%; Pred. No. 7.3e-58;
Matches 133; Conservative 70; Mismatches 132; Indels 7; Gaps 5;

```
Qy         18 QALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSL 77
              || |:  : || :|  ||  |:||: :||  |: |  :| ||| | |   |     :  :
Db         11 QAKLWNFIYGFAESLVLKCAVQLDLANII--HNSGTSMTLSELSSRLPSQPVNEDALYRV 68

Qy         78 MRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQLKR 137
              |||| |   |     |  :  | |   :: ||| :  :   :    : :    |: ||
Db         69 MRYLVHMKLFTKASIDGELR-YGLAPPAKYLVKGWDKCMVGSILAITDKDFMAPWHYLKD 127

Qy        138 WVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAF-RDCNWVFEGLES 196
              :  |  | || :|||  | :: : |  |: |||||| ||:::  |  ::|  :| |: :
Db        128 GLSGESGTAFEKALGTNIWGYMAEHPEKNQLFNEAMANDSRLIMSALVKECGNIFNGITT 187

Qy        197 IVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVL 256
              :|||||||   : |  ||| :|| | : |:|: :  | : :  | ||||| ||||||::
Db        188 LVDVGGGTGTAVRNIANAFPHIKCTVYDLPHVIADSPGYSEVHCVAGDMFKFIPKADAIM 247

Qy        257 LKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM 316
              :| :||:|:| :|::||: ||||:    | |||:::| |:|  : |   |:::| :|: |
Db        248 MKCILHDWDDKECIEILKRCKEAV--PVKGGKVIIVDIVLNV-QSEHPYTKMRLTLDLDM 304

Qy        317 ACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
                  |||| ||:|||| :||:: :||:  |   |:|| ||
Db        305 MLNTGGKERTEEEWKKLIHDAGYKGHKITQITAVQSVIEAYP 346
```

```
RESULT 7
AAY05661
ID   AAY05661 standard; Protein; 375 AA.
XX
AC   AAY05661;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Maize caffeic O-methyltransferase.
XX
KW   Maize; corn; caffeic O-methyltransferase; lignin; transgenic plant.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "encoded by ACC"
XX
PN   WO9910498-A2.
XX
PD   04-MAR-1999.
XX
PF   24-AUG-1998;   98WO-US17519.
XX
PR   12-MAY-1998;   98US-0076851.
PR   27-AUG-1997;   97US-0057082.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bowen BA,  Helentjaris TG,  Wang X;
XX
DR   WPI; 1999-204667/17.
DR   N-PSDB; AAX25200.
XX
PT   Nucleic acids encoding plant lignin biosynthesis enzymes - used to
PT   transform plants to modulate lignin biosynthesis
XX
PS   Claim 9; Page 94-95; 166pp; English.
XX
CC   The present sequence is a caffeic O-methyltransferase, encoded by
CC   a cDNA clone (see AAX25200) isolated from a line A643 cDNA library.
CC   The invention provides methods and compositions relating to
CC   altering lignin biosynthesis and/or the lignin composition of
CC   plants.  Isolated nucleic acids (see AAX25196-216) that code for
CC   proteins (see AAY05657-77) involved in lignin biosynthesis are
CC   claimed.  Also claimed are recombinant expression cassettes, host
CC   cells (especially maize or sorghum), and transgenic plants and
CC   seeds.  The claimed nucleic acids can be used to transform a plant
CC   to modulate lignin biosynthesis.  A claimed method involves
CC   transforming a plant cell with a recombinant expression cassette
CC   comprising a lignin biosynthesis polynucleotide operably linked to
CC   a promoter, growing the plant cell under plant growing conditions,
CC   and inducing expression of the polynucleotide for a time sufficient
CC   to modulate (preferably increase) lignin biosynthesis in the plant.
CC   The plant lignins can be used as chemical feedstock.  Plant
CC   material of increased lignin content can be used as a fuel source,
CC   and in the pulp and paper industry.  Decreased lignin content
CC   improves the digestibility of fodder crops.
XX
SQ   Sequence   375 AA;

  Query Match             31.3%;  Score 590.5;  DB 20;  Length 375;
  Best Local Similarity   37.3%;  Pred. No. 5.7e-53;
  Matches  139;  Conservative   62;  Mismatches  141;  Indels   31;  Gaps   12;

Qy         11 ASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTK 70
              : :: |     |  | | |  |  |   ::| ||| || |  |    |  ::::   : |:|
Db          9 SQDLLQAHDELLHHSLCFAKSLALAVALDLRIPDAIHHHGAG-GATLLQILAETALHPSK 67

Qy         71 TRQVQSLMRYLAHNGFFEIVR----------IH----DNIEAYALTAASELLVKS--SE 113
                | :: ||| |    | | :|           :|    : :  | |||||  ||    |
Db         68 LRALRRLMRVLTVTGIFSVVEQPPAGGGDDSTVHTSDDEAVVVYRLTAASRFLVSDDVST 127

Qy        114 LSLAPMVEYFLEP--NCQGAWNQLKRWVHEEDLTVFEVSLG---TP-FWDFINKDPAYNK 167
               :||| |   |:|  |  |  :  | :|     |    ||  |:  :     |
Db        128 ATLAPFVSLALQPIAACPHALG-ISAWFRQEQHEPSPYGLAFRQTPTIWEHADD---VNA 183

Qy        168 SFNEAMACDSQ-MLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERP 226
               |: || ||: :: :  |:|   | |::|:|||||| |  |  |  ||| ||| ||: |
Db        184 LLNKGMAADSRFLMPIVLRECGETFRGIDSLVDVGGGHGGAAAAIAAAFPHLKCSVLDLP 243

Qy        227 NVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKT 286
               :||       |: || |:||: || | || ||  ||:| |::|:|||:|||:|||
Db        244 HVVAGAPSDGNVQFVAGNMFESIPPATAVFLKKTLHDWGDDECVKILKNCKQAISPRDAG 303

Qy        287 GKVVVIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISP 346
               |||:::| |:   :  :  | ::: |::|   :|| || |::|||:| ||||: ||| |
Db        304 GKVIILDVVVGYKQSNIKHQETQVMFDLYMMA-VNGVERDEQEWKKIFTEAGFKDYKILP 362

Qy        347 FTGYLS-LIEIYP 358
               | :| :||:||
Db        363 VIGDVSVIIEVYP 375


RESULT 8
AAY27183
ID   AAY27183 standard; Protein; 350 AA.
XX
AC   AAY27183;
XX
DT   20-SEP-1999  (first entry)
XX
DE   (S)-3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase (4'-OMT).
XX
KW   (S)-3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase; 4'-OMT;
KW   enzyme; secondary metabolite; reticuline; drug.
XX
OS   Coptis japonika.
XX
PN   JP11178579-A.
XX
PD   06-JUL-1999.
XX
PF   24-DEC-1997;   97JP-0355320.
XX
PR   24-DEC-1997;   97JP-0355320.
XX
PA   (MITC ) MITSUI PETROCHEM IND CO LTD.
XX
DR   WPI; 1999-437312/37.
DR   N-PSDB; AAX89255.
XX
PT   New (S)-3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase  - and
PT   gene encoding it
XX
PS   Claim 1; Fig 12; 28pp; Japanese.
XX
CC   This represents a polypeptide having the enzymatic activity of (S)-3'
CC   -hydroxy-N-methylcoclaurine 4'-O-methyltransferase (4'-OMT). A vector
CC   containing the DNA can be used to transform a microbe for the recombinant
CC   preparation of the enzyme. The enzyme can be used to prepare a secondary
CC   metabolite of a plant derived from reticuline which is useful as a raw
CC   material for drugs.
XX
SQ   Sequence   350 AA;

  Query Match             30.8%;  Score 580.5;  DB 20;  Length 350;
  Best Local Similarity   34.0%;  Pred. No. 5.8e-52;
  Matches  118;  Conservative   83;  Mismatches  133;  Indels   13;  Gaps    6;

Qy         16 QGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQ 75
              : || ::| : || ||  |:  ||| | ||| ::  ||:  ::| | | |       :
Db         13 KAQAHVWKIIYGFADSLVLRCAVELGIVDIIDNNN--QPMALADLASKLPVSDVNCDNLY 70

Qy         76 SLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQL 135
```

```
                 ::|||      : :  |  : |||  : || :::: |: ||:    : :    |: :
Db          71 RILRYLVKMEILRVEKSDDGQKKYALEPIATLLSRNAKRSMVPMILGMTQKDFMTPWHSM 130

Qy         136 KRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAF----RDCNWVF 191
               |  : :     || ::|   |:::   |  :: ||| || ::::|  :      ||   :|
Db         131 KDGLSDNG-HAFEKAMGMTIWEYLEGHPDQSQLFNEGMAGETRLLTSSLISGSRD---MF 186

Qy         192 EGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK 251
               :|::|:||||||  | |  |  |  :|||  :||  :  : |:|: |       |:  :|||||| :|
Db         187 QGIDSLVDVGGGNGTTVKAISDAFPHIKCTLFDLPHVIANSYDLPNIERIGGDMFKSVPS 246

Qy         252 ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLL 311
               | |::|||:||:|||  | :|||: |: |:  :   |||:::|   ::|  |  :::  :|:
Db         247 AQAIILKLILHDWNDEDSIKILKQCRNAVPKDG--GKVIIVDVALDEESD-HELSSTRLI 303

Qy         312 MDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
               :|: |      |||| :| |:|:     |||    ||        |:||::|
Db         304 LDIDMLVNTGGKERTKEVWEKIVKSAGFSGCKIRHIAAIQSVIEVFP 350
```

RESULT 9
AAY05662
ID   AAY05662 standard; Protein; 370 AA.
XX
AC   AAY05662;
XX
DT   19-JUL-1999 (first entry)
XX
DE   Maize caffeic O-methyltransferase.
XX
KW   Maize; corn; caffeic O-methyltransferase; lignin; transgenic plant.
XX
OS   Zea mays.
XX
PN   WO9910498-A2.
XX
PD   04-MAR-1999.
XX
PF   24-AUG-1998; 98WO-US17519.
XX
PR   12-MAY-1998; 98US-0076851.
PR   27-AUG-1997; 97US-0057082.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bowen BA, Helentjaris TG, Wang X;
XX
DR   WPI; 1999-204667/17.
DR   N-PSDB; AAX25201.
XX
PT   Nucleic acids encoding plant lignin biosynthesis enzymes - used to
PT   transform plants to modulate lignin biosynthesis
XX
PS   Claim 9; Page 95-96; 166pp; English.
XX
CC   The present sequence is a caffeic O-methyltransferase of maize,
CC   encoded by a clone (see AAX25201) isolated from a line A632 library.
CC   The invention provides methods and compositions relating to
CC   altering lignin biosynthesis and/or the lignin composition of
CC   plants. Isolated nucleic acids (see AAX25196-216) that code for
CC   proteins (see AAY05657-77) involved in lignin biosynthesis are
CC   claimed. Also claimed are recombinant expression cassettes, host
CC   cells (especially maize or sorghum), and transgenic plants and
CC   seeds. The claimed nucleic acids can be used to transform a plant
CC   to modulate lignin biosynthesis. A claimed method involves
CC   transforming a plant cell with a recombinant expression cassette
CC   comprising a lignin biosynthesis polynucleotide operably linked to
CC   a promoter, growing the plant cell under plant growing conditions,
CC   and inducing expression of the polynucleotide for a time sufficient
CC   to modulate (preferably increase) lignin biosynthesis in the plant.
CC   The plant lignins can be used as chemical feedstock. Plant
CC   material of increased lignin content can be used as a fuel source,
CC   and in the pulp and paper industry. Decreased lignin content
CC   improves the digestibility of fodder crops.
XX
SQ   Sequence 370 AA;

```
  Query Match             30.1%;  Score 567;  DB 20;  Length 370;
  Best Local Similarity   34.7%;  Pred. No. 1.7e-50;
  Matches  137;  Conservative   64;  Mismatches  132;  Indels   62;  Gaps   11;

Qy           1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
               ||     :   : :: |     |  |  |  |  |   ::| ||| || |   |   ::
Db           1 MALMQESSSQDQDMLQAHDELLHHSLCFAKSLALTVALDLRIPDAIHHHGGG--ATLLQI 58

Qy          61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIH------DNIEA---YALTAASELLVK- 110
               ::   : |:|  | :: |||  |   |  |::        |: ||   |||||  ||
Db          59 LAETGLHPSKLRALRRLMRVLTVTGTFS-VQVQQPPAGSDDDEAVVVYRLTAASRFLVSD 117

Qy         111 --SSELSLAPMVEYFLEPNCQG-------AWNQLKR---------------WVHEEDLT 145
                 |:  :|||  |    |:|           || : ::                | |  :|:
Db         118 EVSTATTLAPFVSLALQPIAASPHALGICAWFRQEQHEPSPYGLAFRQTPTLWEHADDV- 176

Qy         146 VFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQ-MLNLAFRDCNWVFEGLESIVDVGGGT 204
                                      |   |:  |  ||: :: :   |  |   :| |: |:|||||
Db         177 -------------------NALLNKGMVADSRFLMPIVLRQCGEMFRGINSLVDVGGGH 216

Qy         205 GITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNW 264
                |  |   |  ||| :||  ||: |:||       |: || |:||: ||  | ||  || ||:|
Db         217 GGAAAAIAAAFPHVKCSVLDLPHVVAGAPSDGNVQFVAGNMFESIPPATAVFLKKTLHDW 276

Qy         265 NDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHMACIINGKE 324
                  |::|:|||:|||:||          |||:::|  |:   :    :   | ::: |::|     :|| |
Db         277 GDDECVKILKNCKQAIPPRDAGGKVIILDVVVGYKQSNIKHQETQVMFDLYMMA-VNGVE 335

Qy         325 RKEEDWKKLFMEAGFQSYKISPFTGYLS-LIEIYP 358
               | |::|||:| ||||: ||| |   | :| :||:||
Db         336 RDEQEWKKIFAEAGFKDYKILPVIGDVSVIIEVYP 370
```

RESULT 10
AAY05663
ID   AAY05663 standard; Protein; 366 AA.
XX
AC   AAY05663;
XX
DT   19-JUL-1999 (first entry)
XX
DE   Maize caffeic O-methyltransferase.
XX
KW   Maize; corn; caffeic O-methyltransferase; lignin; transgenic plant.
XX
OS   Zea mays.
XX
PN   WO9910498-A2.
XX
PD   04-MAR-1999.
XX
PF   24-AUG-1998; 98WO-US17519.
XX
PR   12-MAY-1998; 98US-0076851.
PR   27-AUG-1997; 97US-0057082.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bowen BA, Helentjaris TG, Wang X;
XX
DR   WPI; 1999-204667/17.
DR   N-PSDB; AAX25202.
XX
PT   Nucleic acids encoding plant lignin biosynthesis enzymes - used to
PT   transform plants to modulate lignin biosynthesis
XX
PS   Claim 9; Page 96-97; 166pp; English.
XX
CC   The present sequence is a caffeic O-methyltransferase of maize,

```
CC   encoded by a clone (see AAX25202) isolated from a maize line B73 cDNA
CC   library.  The invention provides methods and compositions relating
CC   to altering lignin biosynthesis and/or the lignin composition of
CC   plants.  Isolated nucleic acids (see AAX25196-216) that code for
CC   proteins (see AAY05657-77) involved in lignin biosynthesis are
CC   claimed.  Also claimed are recombinant expression cassettes, host
CC   cells (especially maize or sorghum), and transgenic plants and
CC   seeds.  The claimed nucleic acids can be used to transform a plant
CC   to modulate lignin biosynthesis.  A claimed method involves
CC   transforming a plant cell with a recombinant expression cassette
CC   comprising a lignin biosynthesis polynucleotide operably linked to
CC   a promoter, growing the plant cell under plant growing conditions,
CC   and inducing expression of the polynucleotide for a time sufficient
CC   to modulate (preferably increase) lignin biosynthesis in the plant.
CC   The plant lignins can be used as chemical feedstock.  Plant
CC   material of increased lignin content can be used as a fuel source,
CC   and in the pulp and paper industry.  Decreased lignin content
CC   improves the digestibility of fodder crops.
XX
SQ   Sequence   366 AA;

  Query Match             28.7%;  Score 540.5;  DB 20;  Length 366;
  Best Local Similarity   33.3%;  Pred. No. 1e-47;
  Matches  128;  Conservative   70;  Mismatches  125;  Indels   61;  Gaps   11;

Qy         11 ASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTK 70
Db          8 SQDLLEAHDELFHHCLCFAKSLALAVAQDLRIPDAIHHHGGG--ATLHQILAEAALHPSK 65

Qy         71 TRQVQSLMRYLAHNGFFEIVRIHDNIEA-------YALTAASELLVKSSE----LSLAPM 119
Db         66 LRALRRLMRVLTVSGVF-TVQYSSTVDASDGADVVYRLTAASRFLVSDSDEAGTASLAPF 124

Qy        120 VEYFL-----EPNCQG--AWNQLKR----------------WVHEEDLTVFEVSLGTPFW 156
Db        125 ANLALHPIAISPHAVGICAWFRQEQHDPSPYGLAFRQIPTIWEHADNV------------ 172

Qy        157 DFINKDPAYNKSFNEAMACDSQ-MLNLAFRDC-NWVFEGLESIVDVGGGTGITAKIICEA 214
Db        173 ---------NALLNKGLLAESRFLMPIVLRECGDEVFRGIDSLVDVGGGHGGAAATIAAA 223

Qy        215 FPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILE 274
Db        224 FPHVKCSVLDLPHVVAGAPSDACVQFVAGNMFHSIPPATAVFFKTTLCDWGDDECIKILK 283

Qy        275 NCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLF 334
Db        284 NCKQAISPRDEGGKVIIMDVVVGYGQSNMKRLETQVMFDLVMMA-VNGVERDEQEWKEMF 342

Qy        335 MEAGFQSYKISPFTGYLSLIEIYP 358
Db        343 IEAGFKDYKIRPVAGLMSVIEVYP 366


RESULT 11
AAG40137
ID   AAG40137 standard; Protein; 368 AA.
XX
AC   AAG40137;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 49761.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
```

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 27.9%; Score 527; DB 21; Length 368;
Best Local Similarity 34.1%; Pred. No. 2.6e-46;
Matches 126; Conservative 70; Mismatches 151; Indels 22; Gaps 10;

```
Qy          6 NNGRKA-SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSIL 64
              :| ||   |  :    ::|:: || |    |  ::| ||: | :|   ||:| :|| | :
Db          3 DNKRKVLDEEAKASLDIWKYVFGFADIAAAKCAIDLKIPEAIENHPSSQPVTLAELSSAV 62

Qy         65 QVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNI-EAYALTAASE--LLVKSSELSLAPMVE 121
               |:  |::   ||:| | | |: :   | :   |  |  |   :: :    |||| |
Db         63 SASPSHLRRI---MRFLVHQGIFKEIPTKDGLATGYVNTPLSRRLMITRRDGKSLAPFVL 119

Qy        122 YFLEPNCQGAWNQLKRW----VHEEDLTVPEVSLGTPFWDFINKDPAYNKSFNEAMACDS 177
              :   |    | :|       |:      |:   |   | |   :|  :   |||||||:
Db        120 FETTPEMLAPWLRLSSVVSSPVNGSTPPPFDAVHGKDVWSFAQDNPFLSDMINEAMACDA 179

Qy        178 -QMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSN 236
               :::      |: :|:|: ::|||||||| |  :: : || :|    : |:|:|     :
Db        180 RRVVPRVAGACHGLFDGVTTMVDVGGGTGETMGMLVKEFPWIKGFNFDLPHVIEVAEVLD 239

Qy        237 NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI 296
              :  | ||||  ||  ||: :| |||:| | ||:|||:|||||:      |||:::::||
Db        240 GVENVEGDMFDSIPACDAIFIKWVLHDWGDKDCIKILKNCKEAV--PPNIGKVLIVESVI 297

Qy        297 NENK-----DER--QVTELKLLMD-VHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
                |||     |||  ::  ::|::| | ||    ||||  ::|  :  ||||  |::
Db        298 GENKKTMIVDERDEKLEHVRLMLDMVMMAHTSTGKERTLKEWDFVLKEAGFARYEVRDID 357

Qy        349 GYLSLIEIY 357
                |||  |
Db        358 DVQSLIIAY 366
```

RESULT 12
AAG40136

ID AAG40136 standard; Protein; 382 AA.
XX
AC AAG40136;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49760.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             27.9%;  Score 527;  DB 21;  Length 382;
  Best Local Similarity   34.1%;  Pred. No. 2.8e-46;
  Matches  126;  Conservative   70;  Mismatches  151;  Indels   22;  Gaps   10;

Qy          6 NNGRKA-SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSIL 64
              :| ||   |  :    ::|:: || |     |  ::| ||: | :|   ||:| :|| | :
Db         17 DNKRKVLDEEAKASLDIWKYVFGFADIAAAKCAIDLKIPEAIENHPSSQPVTLAELSSAV 76

Qy         65 QVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNI-EAYALTAASE--LLVKSSELSLAPMVE 121
                |:  |::   ||:| | | |: :   | :   |  |  |   :: :    |||| |
Db         77 SASPSHLRRI---MRFLVHQGIFKEIPTKDGLATGYVNTPLSRRLMITRRDGKSLAPFVL 133

Qy        122 YFLEPNCQGAWNQLKRW----VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDS 177
              :   |      | :|       |:      |:   |  | |   :|  :   |||||||:
Db        134 FETTPEMLAPWLRLSSVVSSPVNGSTPPPFDAVHGKDVWSFAQDNPFLSDMINEAMACDA 193

Qy        178 -QMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSN 236
               :::      |: :|:|: ::|||||||| |  :: : || :|    : |:|:|     :

Db        194 RRVVPRVAGACHGLFDGVTTMVDVGGGTGETMGMLVKEFPWIKGFNFDLPHVIEVAEVLD 253

Qy        237 NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI 296
               :  | ||||  ||  ||: :| |||:| | ||:|||:|||||:      |||::::||
Db        254 GVENVEGDMFDSIPACDAIFIKWVLHDWGDKDCIKILKNCKEAV--PPNIGKVLIVESVI 311

Qy        297 NENK-----DER--QVTELKLLMD-VHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
              |||      |||  ::  ::|::| | ||    ||||  ::|  :  ||||  |::
Db        312 GENKKTMIVDERDEKLEHVRLMLDMVMMAHTSTGKERTLKEWDFVLKEAGFARYEVRDID 371

Qy        349 GYLSLIEIY 357
               |||  |
Db        372 DVQSLIIAY 380


RESULT 13
ABB93248
ID   ABB93248 standard; Protein; 382 AA.
XX
AC   ABB93248;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 2459.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 2459; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   382 AA;

  Query Match             27.9%;  Score 527;  DB 23;  Length 382;
  Best Local Similarity   34.1%;  Pred. No. 2.8e-46;
  Matches  126;  Conservative   70;  Mismatches  151;  Indels   22;  Gaps   10;

Qy          6 NNGRKA-SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSIL 64
              :| ||   |  :    ::|:: || |     |  ::| ||: | :|   ||:| :|| | :
Db         17 DNKRKVLDEEAKASLDIWKYVFGFADIAAAKCAIDLKIPEAIENHPSSQPVTLAELSSAV 76

Qy         65 QVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNI-EAYALTAASE--LLVKSSELSLAPMVE 121
                |:  |::   ||:| | | |: :   | :   |  |  |   :: :    |||| |
Db         77 SASPSHLRRI---MRFLVHQGIFKEIPTKDGLATGYVNTPLSRRLMITRRDGKSLAPFVL 133
```

```
Qy        122 YFLEPNCQGAWNQLKRW----VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDS 177
              :   |      | :|       |:      |:   |   | |   :|  :   |||||||:
Db        134 FETTPEMLAPWLRLSSVVSSPVNGSTPPPFDAVHGKDVWSFAQDNPFLSDMINEAMACDA 193

Qy        178 -QMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSN 236
               :::     |: :|:|: ::|||||||| |  :: : || :|    : |:|:|     :
Db        194 RRVVPRVAGACHGLFDGVTTMVDVGGGTGETMGMLVKEFPWIKGFNFDLPHVIEVAEVLD 253

Qy        237 NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI 296
              :  | ||||  ||  ||: :| |||:| | ||:|||:|||||:      |||:::::||
Db        254 GVENVEGDMFDSIPACDAIFIKWVLHDWGDKDCIKILKNCKEAV--PPNIGKVLIVESVI 311

Qy        297 NENK-----DER--QVTELKLLMD-VHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
              |||     |||  ::  ::|::| | ||    ||||  ::|  :  ||||  |::
Db        312 GENKKTMIVDERDEKLEHVRLMLDMVMMAHTSTGKERTLKEWDFVLKEAGFARYEVRDID 371

Qy        349 GYLSLIEIY 357
                |||  |
Db        372 DVQSLIIAY 380


RESULT 14
AAG28589
ID   AAG28589 standard; Protein; 368 AA.
XX
AC   AAG28589;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 33864.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
```

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 27.6%; Score 521; DB 21; Length 368;
Best Local Similarity 33.9%; Pred. No. 1.1e-45;
Matches 125; Conservative 70; Mismatches 152; Indels 22; Gaps 10;

```
Qy          6 NNGRKA-SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSIL 64
              :| ||   |  :    ::|:: || |    |  ::| ||: | :|   ||:| :|| | :
Db          3 DNKRKVLDEEAKASLDIWKYVFGFADIAAAKCAIDLKIPEAIENHPSSQPVTLAELSSAV 62

Qy         65 QVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNI-EAYALTAASE--LLVKSSELSLAPMVE 121
               |:  |::   ||:| | | |: :   | :   |  |  |   :: :    |||| |
Db         63 SASPSHLRRI---MRFLVHQGIFKEIPTKDGLATGYVNTPLSRRLMITRRDGKSLAPFVL 119

Qy        122 YFLEPNCQGAWNQLKRW----VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDS 177
              :   |      | :|       |:      |:   |   | |   :|  :   ||| |||:
Db        120 FETTPEMLAPWLRLSSVVSSPVNGSTPPPFDAVHGKDVWSFAQDNPFLSDMINEAXACDA 179

Qy        178 -QMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSN 236
               :::       |: :|:|: ::||||||| |  :: : || :|    : |:|:|     :
Db        180 RRVVPRVAGACHGLFDGVTTMVDVGGGTGETMGMLVKEFPWIKGFNFDLPHVIEVAEVLD 239

Qy        237 NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI 296
              :  | ||||  ||  ||: :| ||:|: | ||:|||:|||||:      |||:::::||
Db        240 GVENVEGDMFDSIPACDAIFIKWVLHDWGDKDCIKILKNCKEAV--PPNIGKVLIVESVI 297

Qy        297 NENK-----DER--QVTELKLLMD-VHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
               |||     |||  ::  ::|::| | ||    ||||  ::|  :  ||||  |::
Db        298 GENKKTMIVDERDEKLEHVRLMLDMVMMAHTSTGKERTLKEWDFVLKEAGFARYEVRDID 357

Qy        349 GYLSLIEIY 357
                 |||  |
Db        358 DVQSLIIAY 366
```

RESULT 15
AAG28588
ID AAG28588 standard; Protein; 382 AA.
XX
AC AAG28588;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33863.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 27.6%; Score 521; DB 21; Length 382;
Best Local Similarity 33.9%; Pred. No. 1.2e-45;
Matches 125; Conservative 70; Mismatches 152; Indels 22; Gaps 10;

```
Qy          6 NNGRKA-SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSIL 64
              :| ||   |  :    ::|:: || |    |  ::| ||: | :|   ||:| :|| | :
Db         17 DNKRKVLDEEAKASLDIWKYVFGFADIAAAKCAIDLKIPEAIENHPSSQPVTLAELSSAV 76

Qy         65 QVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNI-EAYALTAASE--LLVKSSELSLAPMVE 121
                |:  |::   ||:| | | |: :   | :   |  | |   :: :    |||| |
Db         77 SASPSHLRRI---MRFLVHQGIFKEIPTKDGLATGYVNTPLSRRLMITRRDGKSLAPFVL 133

Qy        122 YFLEPNCQGAWNQLKRW----VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDS 177
              :   |      | :|       |:      |:   |  | |   :|  :   ||| |||:
Db        134 FETTPEMLAPWLRLSSVVSSPVNGSTPPPFDAVHGKDVWSFAQDNPFLSDMINEAXACDA 193

Qy        178 -QMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSN 236
               :::      |: :|:|: ::||||||| |  :: : || :|    : |:|:|     :
Db        194 RRVVPRVAGACHGLFDGVTTMVDVGGGTGETMGMLVKEFPWIKGFNFDLPHVIEVAEVLD 253

Qy        237 NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI 296
              :  | ||||  ||  ||: :| |||:| | ||:|||:|||||:      |||:::::||
Db        254 GVENVEGDMFDSIPACDAIFIKWVLHDWGDKDCIKILKNCKEAV--PPNIGKVLIVESVI 311

Qy        297 NENK-----DER--QVTELKLLMD-VHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
               |||     |||  ::  ::|::| | ||    ||||  ::|  :  ||||  |::
Db        312 GENKKTMIVDERDEKLEHVRLMLDMVMMAHTSTGKERTLKEWDFVLKEAGFARYEVRDID 371

Qy        349 GYLSLIEIY 357
               |||  |
Db        372 DVQSLIIAY 380
```

Search completed: November 10, 2003, 03:18:47
Job time : 82 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 10, 2003, 03:17:25 ; Search time 26 Seconds
(without alignments)
582.588 Million cell updates/sec

Title: US-09-868-547-4
Perfect score: 1886
Sequence: 1 MASSLNNGRKASEIFQGQAL.....FQSYKISPFTGYLSLIEIYP 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	426	22.6	363	2	US-08-845-742-2	Sequence 2, Appli
2	390.5	20.7	364	2	US-08-204-288-5	Sequence 5, Appli
3	376.5	20.0	364	2	US-08-204-288-2	Sequence 2, Appli
4	371.5	19.7	365	2	US-08-715-325-2	Sequence 2, Appli
5	367	19.5	371	4	US-09-500-569-16	Sequence 16, Appl
6	364.5	19.3	365	4	US-09-500-569-4	Sequence 4, Appli
7	357	18.9	358	4	US-09-500-569-18	Sequence 18, Appl
8	350.5	18.6	368	3	US-08-991-677-6	Sequence 6, Appli
9	338.5	17.9	354	4	US-09-500-569-12	Sequence 12, Appl
10	337	17.9	365	2	US-08-204-288-7	Sequence 7, Appli
11	329.5	17.5	362	4	US-09-500-569-14	Sequence 14, Appl
12	322.5	17.1	356	4	US-09-500-569-6	Sequence 6, Appli
13	319	16.9	378	4	US-09-500-569-10	Sequence 10, Appl
14	303	16.1	365	1	US-08-186-833-4	Sequence 4, Appli
15	283.5	15.0	328	4	US-09-615-192A-274	Sequence 274, App
16	273.5	14.5	351	4	US-09-266-965-133	Sequence 133, App
17	268	14.2	356	1	US-07-959-941-2	Sequence 2, Appli
18	268	14.2	356	1	US-08-259-924-2	Sequence 2, Appli
19	258.5	13.7	160	4	US-09-615-192A-275	Sequence 275, App
20	255	13.5	174	4	US-09-500-569-26	Sequence 26, Appl
21	227.5	12.1	156	4	US-09-615-192A-272	Sequence 272, App
22	218.5	11.6	305	4	US-09-500-569-2	Sequence 2, Appli
23	214	11.3	188	4	US-09-500-569-8	Sequence 8, Appli
24	202	10.7	145	4	US-09-615-192A-270	Sequence 270, App
25	194	10.3	345	3	US-09-059-522-5	Sequence 5, Appli
26	194	10.3	345	3	US-09-382-027-5	Sequence 5, Appli
27	189.5	10.0	452	3	US-09-059-522-3	Sequence 3, Appli
28	189.5	10.0	452	3	US-09-382-027-3	Sequence 3, Appli
29	189.5	10.0	621	3	US-09-059-522-1	Sequence 1, Appli
30	189.5	10.0	621	3	US-09-382-027-1	Sequence 1, Appli
31	186.5	9.9	115	4	US-09-266-965-3	Sequence 3, Appli
32	135.5	7.2	368	4	US-09-252-991A-20452	Sequence 20452, A
33	134.5	7.1	139	4	US-09-252-991A-20476	Sequence 20476, A
34	122.5	6.5	231	4	US-09-500-569-20	Sequence 20, Appl
35	120.5	6.4	117	4	US-09-266-965-5	Sequence 5, Appli
36	119	6.3	94	1	US-08-266-451B-20	Sequence 20, Appl
37	119	6.3	94	2	US-08-748-725-20	Sequence 20, Appl
38	110.5	5.9	115	4	US-09-266-965-4	Sequence 4, Appli
39	98.5	5.2	1890	4	US-09-004-838-88	Sequence 88, Appl
40	98	5.2	541	2	US-08-540-804-16	Sequence 16, Appl
41	98	5.2	541	2	US-08-218-265-16	Sequence 16, Appl
42	98	5.2	541	3	US-08-521-872-16	Sequence 16, Appl
43	98	5.2	541	3	US-08-590-399-16	Sequence 16, Appl
44	97.5	5.2	983	2	US-08-449-645A-21	Sequence 21, Appl
45	97.5	5.2	983	2	US-08-702-367A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-845-742-2
; Sequence 2, Application US/08845742C
; Patent No. 5973229
; GENERAL INFORMATION:
; APPLICANT: Walton, Jonathan D
; APPLICANT: Scott-Craig, John S
; TITLE OF INVENTION: Gene Encoding Herbicide Safener Binding Protein
; FILE REFERENCE: 6550-000007
; CURRENT APPLICATION NUMBER: US/08/845,742C
; CURRENT FILING DATE: 1997-04-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Zea mays
US-08-845-742-2

Query Match 22.6%; Score 426; DB 2; Length 363;
Best Local Similarity 31.4%; Pred. No. 9.2e-36;
Matches 116; Conservative 70; Mismatches 144; Indels 40; Gaps 10;

```
Qy          12 SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKT 71
               :|: : || ::: | ::    |:  ||| ||  |  : ||   : :|||: | :| ||
Db          11 AELLKAQADIWRLSLSYLTPLSLRCAVELGIPTAI--YRHGGAASAAELVTALSLPSTKL 68

Qy          72 RQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLV---------KSSELSLAPMVEY 122
                 :: |:| || :| | : :     | | ::  | |||           : | |
Db          69 PFLRRLLRLLAASGVFTVDKQSSEEERYRISPVSYLLVDGIPHEDHMNHTALVLTCTSTR 128

Qy         123 FLEPNCQGAWNQLKRWVHE-------EDL---TVFEVSLGTPFWDFINKDPAYNKSFNEA 172
               ::|     |   | |           |:|   |:|  |:|       : |  ::   :||
Db         129 YIE-----AGIGLAEWFKRDVVTSPPEELHGATLFHESMG-------SLDADFHDMASEA 176

Qy         173 M-ACDSQMLNLAFRDCNWVFEGLESIVDVGGGTGIT--AKIICEAFPKLKCMVLERPNVV 229
               : | |:  : :| |:  :|||::|:    |  |    |: | :||| :|| ||  | ::
Db         177 LDAHDNFGIEIAMREFRDLFEGIQSMTYCCGNFGDDKGARAIVKAFPHIKCTVLAPPKII 236

Qy         230 ENLSGSN-NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGK 288
                   : :| |||  || |  |:||||||:  | :|:|:|  |::||      ||
Db         237 ATKPADGAMINYVEGDMFSFIPPAQTVVLKLVLHHLTDEECVKLLAQCRKAIPSRKDGGK 296

Qy         289 VVVIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
               |:: | :|: :     : |  ||||: |  :  |::| |::| :|| :|||  |||
Db         297 VIIGDIMIDHS---GPMLETHLLMDIGMMTMTKGRQRDEKEWSELFTKAGFSEYKILKEF 353

Qy         349 GYLSLIEIYP 358
               |     |:||
Db         354 GARVAFEVYP 363
```

```
RESULT 2
US-08-204-288-5
; Sequence 5, Application US/08204288
; Patent No. 5959178
;  GENERAL INFORMATION:
;    APPLICANT:  VAN DOORSSELAERE, Jan
;    APPLICANT:  FRITIG, Bernard J.M.
;    APPLICANT:  INZE, Dirk G.
;    APPLICANT:  JOUANIN, Lise
;    APPLICANT:  KNIGHT, Mary E.
;    APPLICANT:  VAN MONTAGU, Marc
;    APPLICANT:  LEGRAND, Michel
;    TITLE OF INVENTION:  MODIFICATION OF LIGNIN SYNTHESIS IN
;    TITLE OF INVENTION:  PLANTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,288
;      FILING DATE:  10-MAR-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9119279.9
;      FILING DATE:  10-SEP-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/01460
;      FILING DATE:  09-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOKULIS, Paul N.
;      REGISTRATION NUMBER:  16,773
;      REFERENCE/DOCKET NUMBER:  206860/SEE36543/UST
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  364 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-204-288-5

  Query Match            20.7%;  Score 390.5;  DB 2;  Length 364;
  Best Local Similarity  30.4%;  Pred. No. 4.4e-32;
  Matches  103;  Conservative   64;  Mismatches  135;  Indels   37;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL-----QVPPTKTRQVQSLMRYLAHNGFFE 88
              ||  ||||: :::     |  |: ||| : |     : | |  | :: | |  |   |
Db         36 LKSAVELDLLELMAKAGPGAAISPSELAAQLSTQNPEAPVMLDRMLRLLASYSVLNCTLR 95

Qy         89 IVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHEED 143
                 :       |:|    : | |::: :|:||::            ||    :: | | :|
Db         96 TLPDSSVERLYSLAPVCKYLTKNADGVSVAPLL----------LMNQDKVLMESWYHLKD 145

Qy        144 LTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
               :     |  : |   :::   || :|| ||  |:  | |      :    |||| |||
Db        146 AVLDGGIPFNKAYGMTAFEYHGTDPRFNKVFNRGMSDHSTMSMKKILEDYKGFEGLNSIV 205

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
              |||||||  |  :|   :| :| :  : |:|: :      :  |||||  :|||||: :|
Db        206 DVGGGTGATVNMIVSKYPSIKGINFDLPHVIGDAPTYPGVEHVGGDMFASVPKADAIFMK 265

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
               : |:|:|  |:| |:|| ||:      |||:: : :: |  |    |:  : :|: | |
Db        266 WICHDWSDEHCLKFLKNCYEALPA---NGKVIIAECILPEAPDTSLATKNTVHVDIVMLA 322

Qy        318 CIINGKERKEEDWKKLFMEAGFQSYKISPPTGYLSLIEI 356
                 |||| |:::: |   ||        |||:  |: :
Db        323 HNPGGKERTEKEFEALAKGAG--------FTGFARLVAL 353


RESULT 3
US-08-204-288-2
; Sequence 2, Application US/08204288
; Patent No. 5959178
;  GENERAL INFORMATION:
;    APPLICANT:  VAN DOORSSELAERE, Jan
;    APPLICANT:  FRITIG, Bernard J.M.
;    APPLICANT:  INZE, Dirk G.
;    APPLICANT:  JOUANIN, Lise
;    APPLICANT:  KNIGHT, Mary E.
;    APPLICANT:  VAN MONTAGU, Marc
;    APPLICANT:  LEGRAND, Michel
;    TITLE OF INVENTION:  MODIFICATION OF LIGNIN SYNTHESIS IN
;    TITLE OF INVENTION:  PLANTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,288
;      FILING DATE:  10-MAR-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9119279.9
;      FILING DATE:  10-SEP-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/01460
;      FILING DATE:  09-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOKULIS, Paul N.
;      REGISTRATION NUMBER:  16,773
;      REFERENCE/DOCKET NUMBER:  206860/SEE36543/UST
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  364 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-204-288-2

  Query Match            20.0%;  Score 376.5;  DB 2;  Length 364;
  Best Local Similarity  30.4%;  Pred. No. 1.3e-30;
  Matches   98;  Conservative   69;  Mismatches  136;  Indels   19;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFFEI 89
              ||  :|||: :|:     |  :: ||: | |  |||       :  ::| ||      |
Db         36 LKTAIELDLLEIMAKAGPGAFLSTSEIASHL---PTKNPDAPVMLDRILRLLASYSIL-I 91
```

```
Qy          90 VRIHDNIEA-----YALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRWVHEED 143
Db          92 CSLKDHPDGKVERLYGLAPVCKFLTKNEDGVSVSPLCLMNQDKVLMESWYYLKDAILDGG 151

Qy         144 LTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGG 203
Db         152 IP-FNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYKGFEGLTSLVDVGGG 210

Qy         204 TGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHN 263
Db         211 TGAVVNTIVSKYPSIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPKADAVFMKWICHD 270

Qy         264 WNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-ACIING 322
Db         271 WSDAHCLKFLKNCYDAL---PENGKVILVECILPVAPDTSLATKGVVHIDVIMLAHNPGG 327

Qy         323 KERKEEDWKKLFMEAGFQSYKI 344
Db         328 KERTEKEFEGLAKGAGFQGFEV 349
```

RESULT 4
US-08-715-325-2
; Sequence 2, Application US/08715325
; Patent No. 5886243
; GENERAL INFORMATION:
; APPLICANT: Chiang, Vincent L
; APPLICANT: Tsai, Chung-Jui
; APPLICANT: Podila, Gopi
; TITLE OF INVENTION: Genetic Engineering of Wood Color
; TITLE OF INVENTION: in Plants
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael, Best & Friedrich
; STREET: 100 E. Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: United States of America
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS 6.22
; SOFTWARE: WordPerfect for Windows 5.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,325
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/007727
; FILING DATE: 30 NOV 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gemignani, Joseph A
; REGISTRATION NUMBER: 19,482
; REFERENCE/DOCKET NUMBER: 66040/9627
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 271-6560
; TELEFAX: (414) 277-0656
; TELEX: 262057
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; HYPOTHETICAL: no
US-08-715-325-2

Query Match 19.7%; Score 371.5; DB 2; Length 365;
Best Local Similarity 29.7%; Pred. No. 4.2e-30;
Matches 97; Conservative 69; Mismatches 132; Indels 29; Gaps 8;

```
Qy          34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFF-- 87
Db          36 LKTAIELDLLEIMAKAGPGAFLSTSEIASHL---PTKNPDAPVMLDRILRLLASYSILTC 92

Qy          88 --------EIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRW 138
Db          93 SLKDLPDGKVERL------YGLAPVCKFLTKNEDGVSVSPLCLMNQDKVLMESWYYLKDA 146

Qy         139 VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
Db         147 ILDGGIP-FNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYKGFEGLTSLV 205

Qy         199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
Db         206 DVGGGTGAVVNTIVSKYPSIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPKADAVFMK 265

Qy         259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
Db         266 WICHDWSDAHCLKFLKNCYDAL---PENGKVILVECILPVAPDTSLATKGVVHVDVIMLA 322

Qy         318 CIINGKERKEEDWKKLFMEAGFQSYKI 344
Db         323 HNPGGKERTEKEFEGLAKGAGFQGFEV 349
```

RESULT 5
US-09-500-569-16
; Sequence 16, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
; FILE REFERENCE: BB1327 US NA
; CURRENT APPLICATION NUMBER: US/09/500,569
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,587
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-500-569-16

Query Match 19.5%; Score 367; DB 4; Length 371;
Best Local Similarity 29.7%; Pred. No. 1.3e-29;
Matches 104; Conservative 70; Mismatches 142; Indels 34; Gaps 10;

```
Qy          27 GFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQ----VQSLMRYLA 82
Db          34 GFTVSMTLKAAIQLGLIDALTAAADGRALTAGELVA--QLPAVDDAEAATSVDRMLRLLA 91

Qy          83 HNGFFEIVRIH-------DNIEAYA-------LTAASELLVKSSELSLAPMVEYFLEPNC 128
Db          92 S---FNVVRCSTEAGPGGDPLRRYSPAPVCRWFTAGD-----NHQGSLAPRLMLDVDEDN 143

Qy         129 QGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCN 188
Db         144 LSTWHQMAAAVVSGGPSAFERAHGMPLFEYMGTNHRFNMLFNQAMSQQSMMVMNKLLDRF 203

Qy         189 WVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKC 248
Db         204 HGFDGISVLVDVGGGTGVTLKMIISRYKHITGVNFDLPHVISQAPSLPGVNHVAGNMFES 263

Qy         249 IPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTEL 308
Db         264 VPKGDAIFLKSMLLR-NDEECIKILKNCHYAL---SDNGKVIVVDIVLPETPKPVPEAQN 319

Qy         309 KLLMDVHMA-CIINGKERKEEDWKKLFMEAGFQ-SYKISPFTGYLSLIEI 356
```

```
              | |||  |      :   ||  |  |:::  ||  |::||   |:: :          ||:
Db        320 PLRMDVMMLNNLRGGKIRTEQEYAKLAMDSGFSGSFRTTYIFANFMAIEL 369

RESULT 6
US-09-500-569-4
; Sequence 4, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Shen, Jennie
;  TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
;  FILE REFERENCE: BB1327 US NA
;  CURRENT APPLICATION NUMBER: US/09/500,569
;  CURRENT FILING DATE: 2000-02-09
;  EARLIER APPLICATION NUMBER: 60/119,587
;  EARLIER FILING DATE: 1999-February-10
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
;   LENGTH: 365
;   TYPE: PRT
;   ORGANISM: Glycine max
US-09-500-569-4

  Query Match             19.3%;  Score 364.5;  DB 4;  Length 365;
  Best Local Similarity   31.2%;  Pred. No. 2.2e-29;
  Matches  100;  Conservative   67;  Mismatches  138;  Indels   15;  Gaps    9;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVP---PTKTRQVQSLMRYLAHNGF--FE 88
              ||   :|||: :||       |   :: :::  |   |:|    |      :  ::| ||      |
Db         36 LKSALELDLLEIIAKAGPGVHLSPTDISS--QLPTQNPDAPVMLDRILRLLACYNILSFS 93

Qy         89 IVRIHD-NIE-AYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRWVHEEDLT 145
               :  :  |  :|  | |    :: |||: : :|:|  :      :       :|  ||  | |  :
Db         94 LRTLPDGKVERLYGLAPVAKYLVKNEDGVSIAALNLMNQDKVLMESWYYLKDAVLEGGIP 153

Qy        146 VFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGGTG 205
               |  :  |     :::     || :|| ||: ||  |  :       :      || |:|:||||||||
Db        154 -FNKAYGMTAFEYHGTDPRFNKVFNKGMADHSTITMKKILETYTGFESLKSLVDVGGGTG 212

Qy        206 ITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWN 265
                 :|    | :| :  :  |:|:|:       :     |||||  :|||||: :| : |:|:
Db        213 AVINMIVSKHPTIKGINFDLPHVIEDAPSYPGVEHVGGDMFASVPKADAIFMKWICHDWS 272

Qy        266 DNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-ACIINGKE 324
               |  |:| |:|| ||:       |||:| : ::       |         |:  : :|| |    |||
Db        273 DEHCLKFLKNCYEAL---PDNGKVIVAECILPVAPDSSLATKGVVHIDVIMLAHNPRGKE 329

Qy        325 RKEEDWKKLFMEAGFQSYKI 344
               | |::::  |     :|||  :::
Db        330 RTEKEFEALAKGSGFQGFRV 349

RESULT 7
US-09-500-569-18
; Sequence 18, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Shen, Jennie
;  TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
;  FILE REFERENCE: BB1327 US NA
;  CURRENT APPLICATION NUMBER: US/09/500,569
;  CURRENT FILING DATE: 2000-02-09
;  EARLIER APPLICATION NUMBER: 60/119,587
;  EARLIER FILING DATE: 1999-February-10
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
;   LENGTH: 358
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (134)
US-09-500-569-18

  Query Match             18.9%;  Score 357;  DB 4;  Length 358;
  Best Local Similarity   30.7%;  Pred. No. 1.3e-28;
  Matches   99;  Conservative   65;  Mismatches  140;  Indels   18;  Gaps    9;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQV-PPTKTRQVQSLMRYLA-HNGFFEIVR 91
              ||   :|||: :||       |   :: |:: | |      |       :   ::| || :|      :|
Db         36 LKSALELDLLEIIAKAGPGVHLSPSDIASRLPTHNPDAPVMLDRILRLLACYNILSFSLR 95

Qy         92 I--HDNIE-AYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRWVHEEDLTVF 147
                  |  :|  |  |    :: ||:: : :|:|  :       :       :|  ||  |  |  :  |
Db         96 TLPHGKVERLYGLAPVAKYLVRNEDGVSIAALNLMNQDXILMESWYYLKDAVLEGGIP-F 154

Qy        148 EVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGGTGIT 207
                 : |     :::     || :|| ||: ||  |  :       :      ||||:|:||||||||
Db        155 NKAYGMTAFEYHGTDPRFNKVFNKGMADHSTITMKKILETYTGFEGLKSLVDVGGGTGAV 214

Qy        208 AKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDN 267
                :|    :| :|  :  : |:|:|:       :     |||||  :|||||: :| : |:|:|
Db        215 VNMIVSKYPTIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPKADAIFMKWICHDWSDE 274

Qy        268 DCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHMACIINGKERKE 327
                |:| |:|| ||:       |||:| : ::       |         |:  : :|| |     | ||:
Db        275 HCLKFLKNCYEAL---PDNGKVIVAECILPVAPDSSLATKGVVHIDVIMLAHNPGGEREN 331

Qy        328 ED-----WKKLFMEAGFQSYKI 344
                :         | |     :|||  :::
Db        332 KRKSLRLWAK---GSGFQGFQV 350

RESULT 8
US-08-991-677-6
; Sequence 6, Application US/08991677A
; Patent No. 6252135
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent L
;  APPLICANT: Carraway, Daniel T
;  APPLICANT: Smeltzer, Richard H
;  TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
;  FILE REFERENCE: 50617
;  CURRENT APPLICATION NUMBER: US/08/991,677A
;  CURRENT FILING DATE: 1997-12-16
;  EARLIER APPLICATION NUMBER: US 60/033,381
;  EARLIER FILING DATE: 1996-12-16
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 368
;   TYPE: PRT
;   ORGANISM: Liquidambar styraciflua
US-08-991-677-6

  Query Match             18.6%;  Score 350.5;  DB 3;  Length 368;
  Best Local Similarity   29.0%;  Pred. No. 6.4e-28;
  Matches  100;  Conservative   62;  Mismatches  130;  Indels   53;  Gaps   11;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFFEI 89
              ||   :|||: :|:       |   |: |:: |  |     |||         :   ::| ||
Db         38 LKSAIELDVLEIMAKAGPGAHISTSDIASKL---PTKNPDAAVMLDRMLRLLA------- 87

Qy         90 VRIHDNIEAYALTAASELLVKSSEL----SLAPMVEYFLEPNCQGA-------WNQ---- 134
                      :|::    |    :    ::      |||: :  ||   |  |         ||
Db         88 --------SYSVLTCSLRTLPDGKIERLYGLAPVCK-FLTRNDDGVSIAALSLMNQDKVL 138

Qy        135 LKRWVHEEDLTVFEVSLGTPF--------WDFINKDPAYNKSFNEAMACDSQMLNLAFRD 186
              ::  |  |     ||     :   |  ||         :::      || :|   ||  |:   |  :     :
Db        139 MESWYH---LTEAVLEGGIPFNKAYGMTAFEYHGTDPRFNTVFNNGMSNHSTITMKKILE 195
```

```
Qy        187 CNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMF 246
Db        196 TYKGFEGLGSVVDVGGGTGAHLNMIIAKYPMIKGINFDLPHVIEEAPSYPGVEHVGGDMF 255

Qy        247 KCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVT 306
Db        256 VSVPKGDAIFMKWICHDWSDEHCLKFLKKCYEAL---PTNGKVILAECILPVAPDASLPT 312

Qy        307 ELKLLMDVHM-ACIINGKERKEEDWKKLFMEAGFQSYKISPFTGY 350
Db        313 KAVVHIDVIMLAHNPGGKERTEKEFEALAKGAGFEGFRVVASCAY 357
```

RESULT 9
US-09-500-569-12
; Sequence 12, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
; FILE REFERENCE: BB1327 US NA
; CURRENT APPLICATION NUMBER: US/09/500,569
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,587
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Glycine max
US-09-500-569-12

Query Match 17.9%; Score 338.5; DB 4; Length 354;
Best Local Similarity 26.4%; Pred. No. 1.1e-26;
Matches 89; Conservative 74; Mismatches 147; Indels 27; Gaps 8;

```
Qy         38 VELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQ----VQSLMRYLAHNGFFEIVRIH 93
Db         27 IELGIFDIIAKAGEGAKLSAKDIAAKL---PCKNSEGATMLDRILRLLVCHSIIDCTVVA 83

Qy         94 DNIEA--------YALTAASELLVK-SSELSLAPMVEYFLEPNCQGAWNQLKRWVHEEDL 144
Db         84 DQQHGPPPHLQRFYAMNPVAKYFASIDGAGSLGPLMVLTQDKALLHSWYQLKDAILEGGI 143

Qy        145 TVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGGT 204
Db        144 PFNRVH-GKHVFEYSDMNSSFNQLFMAAMTNRATLIMKKIVESYKGFEHLNSLVDVGGGL 202

Qy        205 GITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNW 264
Db        203 GVTLNIVTSKYPHIKGINFDLPHVIEHASTYPGVEHVGGDMFESVPQGDAILMMCVLHDW 262

Qy        265 NDNDCMKILENCKEAISGESKTGKVVVIDTVI----NENKDERQVTELKLLMDVHMACII 320
Db        263 SDEWCLKVLKNCYASIPSD---GKVIVVDGILPFEPKTTGASKSISQFDVLM---MTTNP 316

Qy        321 NGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIY 357
Db        317 GGKERSEEEFMALAKGAGYSGIRFTCFVSDLWVMEFF 353
```

RESULT 10
US-08-204-288-7
; Sequence 7, Application US/08204288
; Patent No. 5959178
; GENERAL INFORMATION:
; APPLICANT: VAN DOORSSELAERE, Jan
; APPLICANT: FRITIG, Bernard J.M.
; APPLICANT: INZE, Dirk G.
; APPLICANT: JOUANIN, Lise
; APPLICANT: KNIGHT, Mary E.
; APPLICANT: VAN MONTAGU, Marc
; APPLICANT: LEGRAND, Michel
; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,288
; FILING DATE: 10-MAR-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9119279.9
; FILING DATE: 10-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01460
; FILING DATE: 09-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 206860/SEE36543/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-204-288-7

Query Match 17.9%; Score 337; DB 2; Length 365;
Best Local Similarity 27.5%; Pred. No. 1.6e-26;
Matches 96; Conservative 76; Mismatches 131; Indels 46; Gaps 12;

```
Qy         20 LLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVP----PTKTRQVQ 75
Db         25 LLSSSVLPFV----LHSTIQLEVFEIL-AKSNDTKLSASQIVS--QIPNCKNPDAATMLD 77

Qy         76 SLMRYLAHNGFF--EIVRIHDN----IEAYALTAASELLVKSSE-LSLAPMVEYFLEPNC 128
Db         78 RMLYVLASYSLFTCSIVEDEENNGGQKRVYGLSQVGKFFVRDEDGASMGPLLALLQDKVF 137

Qy        129 QGAWNQLKRWVHEEDLTVFEVSLGTPF---------WDFINKDPAYNKSFNEAMACDSQM 179
Db        138 INSWFELKDAVLEG---------GVPFDRVHGVVHAFEYPKSDPKFNDVFNKAMINHTTV 188

Qy        180 LNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLT 239
Db        189 VMKKILENYKGFENLKTLVDVGGGLGVNLKMITSKYPTIKGTNFDLPHVVQHAPSYPGVE 248

Qy        240 FVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINEN 299
Db        249 HVGGDMPFSVPEGDAIFMKWILHDWSDSHNLKLLKNCYKAL---PDNGKVIVVEAILPVK 305

Qy        300 KDERQ----VTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKI 344
Db        306 PDIDTAVVGVSQCDLIM---MAQNPGGKERSEEEFRALATEAGFKGVNL 351
```

RESULT 11
US-09-500-569-14
; Sequence 14, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
; FILE REFERENCE: BB1327 US NA
; CURRENT APPLICATION NUMBER: US/09/500,569
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,587
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-500-569-14

Query Match 17.5%; Score 329.5; DB 4; Length 362;
Best Local Similarity 28.0%; Pred. No. 9.4e-26;
Matches 94; Conservative 71; Mismatches 126; Indels 45; Gaps 11;

Qy 34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVP----PTKTRQVQSLMRYLAHNGFFEI 89
Db 35 LKNAIELGMLEILVG-AGGKMLSPSEVAA--QLPSKANPEAPVMVDRMLRLLASN----- 86

Qy 90 VRIHDNIEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWN----------QLKRW 138
Db 87 -----NVVSCEVEEGKDGLL-ARRYGPAPVCKW-LTPNEDGASMAGLLLMTHDKVTMESW 139

Qy 139 VHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEG 193
Db 140 YYLKDVALEGGQPFHRAHGMTAYEYNSTDPRANCLFNEAMLNHSTIITKKLLEFYRGFDN 199

Qy 194 LESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKAD 253
Db 200 VETLVDVAGGVGATAHAITSKYPHIKGVNFDLPHVISEAPPYPGVQHIAGDMFKKVPSGD 259

Qy 254 AVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVT----ELK 309
Db 260 AILLKWILHNWTDDYCMTLLRNCYDAL---PMNGKVVIVEGILPVKPDAMPSTQTMFQVD 316

Qy 310 LLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKIS 345
Db 317 MMMLLHTA---GGKERELSEFEELAKGAGFSTVKTS 349

RESULT 12
US-09-500-569-6
; Sequence 6, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
; FILE REFERENCE: BB1327 US NA
; CURRENT APPLICATION NUMBER: US/09/500,569
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,587
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-500-569-6

Query Match 17.1%; Score 322.5; DB 4; Length 356;
Best Local Similarity 28.0%; Pred. No. 4.9e-25;
Matches 94; Conservative 68; Mismatches 121; Indels 53; Gaps 12;

Qy 34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL--QVPPTKTRQVQSLMRYLAHNGFFEIVR 91
Db 33 LKNAIELGLLETLVA-AGGKLLTPAEVAAKLPSTANPAAADMVDRMLRLLA--------- 82

Qy 92 IHDNIEAYALTAASELLVKSSELS----LAPMVEYFLEPNCQGA-------WNQ----LK 136
Db 83 ------SYNVVSCTMEEGKDGRLSRRYGAAPVCK-FLTPNEDGVSMAALALMNQDKVLME 135

Qy 137 RWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVF 191
Db 136 SWYYLKDAVLDGGIPFNKAYGMSAFEYHGTDPRFNRVFNEGMKNHSIIITKKLLEVYKGF 195

Qy 192 EGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK 251
Db 196 EGLGTIVDVGGGVG----AIVAAYPAIKGINFDLPHVISEAPPFPGVTHVGGDMFQKVPS 251

Qy 252 ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD----ERQVTE 307
Db 252 GDAILMKWILHDWSDEHCATLLKNCYDALPAH---GKVVLVECILPVNPEATPKAQGVFH 308

Qy 308 LKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYK 343
Db 309 VDMIMLAHNP---GGRERYERDFEALAKGAGFKAIK 341

RESULT 13
US-09-500-569-10
; Sequence 10, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
; FILE REFERENCE: BB1327 US NA
; CURRENT APPLICATION NUMBER: US/09/500,569
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,587
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-500-569-10

Query Match 16.9%; Score 319; DB 4; Length 378;
Best Local Similarity 28.9%; Pred. No. 1.2e-24;
Matches 94; Conservative 59; Mismatches 146; Indels 26; Gaps 8;

Qy 34 LKWMVELDIPDIIHSHSHGQPITFSELVS--ILQVPPTKTRQVQSLMRYLAHNGFFEIVR 91
Db 41 LKAAIELGLLDALAAAGDGRALTADELAAARLPDAAPDKAEAASSVDRMLRLLASFDVVK 100

Qy 92 IHDNIEA----------YALTAASELLV---KSSELSLAPMVEYFLEPNCQGAWNQLKRW 138
Db 101 C--STEAGPGGEPPRRRYSPAPVCRLFTAGGNSHRGSLAPSVLFGVDEDYLCTWRQLAAA 158

Qy 139 VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
Db 159 VGGGGPSAFERAHGMRMFEYMGTNRRLNTLFNQAMAQQSMIVIDKLLDRFHGFDGVGVLV 218

Qy 199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK-ADAVLL 257
Db 219 DVGGGTGATLEMITSRYKHITGVNFDLPHVISQAPSIPGVKHIAGNMFESISNIGDAIFL 278

```
Qy        258 KLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHMA 317
Db        279 KMILHMQNDEDCIKILKNCHQAL---PDNGKVIAVEIVLPTIPDLAQTARYPFQMD--MI 333

Qy        318 CIIN---GKERKEEDWKKLFMEAGF 339
Db        334 MLSNSRGGKERTELEFAKLATDSGF 358
```

RESULT 14
US-08-186-833-4
; Sequence 4, Application US/08186833
; Patent No. 5563324
; GENERAL INFORMATION:
; APPLICANT: Tarczynski, Mitchell C.
; APPLICANT: Jensen, Richard G.
; APPLICANT: Bohnert, Hans J.
; TITLE OF INVENTION: Transgenic Plants With Enhanced Mannitol
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: P.O. Box 2113, First Wisconsin Plaza
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word, Version #5.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,833
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/871,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 9221490026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608)251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-186-833-4

Query Match 16.1%; Score 303; DB 1; Length 365;
Best Local Similarity 26.5%; Pred. No. 5.4e-23;
Matches 89; Conservative 61; Mismatches 138; Indels 48; Gaps 9;

```
Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVP-PTKTRQVQSLMRYLAHNGFFEIVRI 92
Db         38 LKSAFELKILDIFSKAGEGVFVSTSEIASQIGAKNPNAPVLLDRMLRLLASHS------- 90

Qy         93 HDNIEAYALTAASELLVKSSE--LSLAPMVEYFLEPNCQGAWNQL----------KRWVH 140
Db         91 -------VLTCKLQKGEGGSQRVYGPAPLCNYLASNDGQGSLGPLLVLHHDKVMMESWFH 143

Qy        141 EEDLTVFEVSLGTPF--------WDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFE 192
Db        144 LND---YILEGGVPFKRAHGMIQFDYTGTDERFNHVFNQGMAHHTILVMKKLLDNYNGFN 200

Qy        193 GLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKA 252
Db        201 DVKVLVDVGGNIGVNVSMIVAKHTHIKGINYDLPHVIADAPSYPGVEHVGGNMFESIPQA 260

Qy        253 DAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI----NENKDERQVTEL 308
Db        261 DAIFMKWVLHDWSDEHCVKILNKCYESL---AKGGKIILVESLIPVIPEDNLESHMVFSL 317

Qy        309 KLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKI 344
Db        318 DCHTLVHNQ---GGKERSKEDFEALASKTGFSTVDV 350
```

RESULT 15
US-09-615-192A-274
; Sequence 274, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 274
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-615-192A-274

Query Match 15.0%; Score 283.5; DB 4; Length 328;
Best Local Similarity 28.0%; Pred. No. 4.8e-21;
Matches 79; Conservative 50; Mismatches 142; Indels 11; Gaps 4;

```
Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVP-PTKTRQVQSLMRYLAHNGFFEIVRI 92
Db         53 LKGVHELGILELL---AKGDQLSPLDIVARLSIDNPAAPDTIDRMLRLLASYSILSCTLV 109

Qy         93 HDN---IEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQLKRWVHEEDLTVFE 148
Db        110 EDKEGRPQRLYGLGPRSKFFLDQNGASTLPTHMLLQEKTLLECWNCLKDAVKEGGADPFT 169

Qy        149 VSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGGTGITA 208
Db        170 RRHGMNVFDYMGQDPRFNDLYNKSMRTGSAIYMPKIAQHYRGFSKAKTVVNVGGGIGETL 229

Qy        209 KIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDND 268
Db        230 KTILSKNPHIRAINYDLPHVIATAPPIPGITHVGGDILKSVPKADVHFLKSVLHRGDDEF 289

Qy        269 CMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKL 310
Db        290 CVKVLKNCWEAL---PPTGKVVIVEEVTPEYPGTDDVSQTTL 328
```

Search completed: November 10, 2003, 03:22:41
Job time : 27 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 10, 2003, 03:21:21 ; Search time 76 Seconds
(without alignments)
809.030 Million cell updates/sec

Title: US-09-868-547-4
Perfect score: 1886
Sequence: 1 MASSLNNGRKASEIFQGQAL.....FQSYKISPFTGYLSLIEIYP 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	590.5	31.3	375	12	US-10-361-460-5	Sequence 5, Appli
2	567	30.1	370	12	US-10-361-460-6	Sequence 6, Appli
3	540.5	28.7	366	12	US-10-361-460-7	Sequence 7, Appli
4	435	23.1	354	12	US-10-361-460-4	Sequence 4, Appli
5	371.5	19.7	365	10	US-09-947-027-6	Sequence 6, Appli
6	371.5	19.7	365	14	US-10-091-009-6	Sequence 6, Appli
7	350.5	18.6	368	9	US-09-796-256A-6	Sequence 6, Appli
8	337.5	17.9	360	12	US-10-289-757-103	Sequence 103, App
9	331.5	17.6	360	12	US-10-289-757-102	Sequence 102, App
10	331.5	17.6	360	12	US-10-289-757-182	Sequence 182, App
11	329.5	17.5	361	12	US-10-289-757-105	Sequence 105, App
12	328.5	17.4	360	12	US-10-289-757-104	Sequence 104, App
13	293	15.5	375	15	US-10-213-473-29	Sequence 29, Appl
14	283.5	15.0	328	16	US-10-174-693-274	Sequence 274, App
15	273.5	14.5	351	12	US-09-953-348-133	Sequence 133, App
16	273.5	14.5	351	15	US-10-267-255-133	Sequence 133, App
17	258.5	13.7	160	16	US-10-174-693-275	Sequence 275, App
18	227.5	12.1	156	16	US-10-174-693-272	Sequence 272, App
19	225.5	12.0	331	15	US-10-156-761-10380	Sequence 10380, A
20	213.5	11.3	285	12	US-09-769-734-15	Sequence 15, Appl
21	202	10.7	145	16	US-10-174-693-270	Sequence 270, App
22	190.5	10.1	359	15	US-10-156-761-9919	Sequence 9919, Ap
23	186.5	9.9	115	12	US-09-953-348-3	Sequence 3, Appli
24	186.5	9.9	115	15	US-10-267-255-3	Sequence 3, Appli
25	169.5	9.0	264	15	US-10-137-036-78	Sequence 78, Appl
26	121.5	6.4	133	15	US-10-267-849-38	Sequence 38, Appl
27	120.5	6.4	117	12	US-09-953-348-5	Sequence 5, Appli
28	120.5	6.4	117	15	US-10-267-255-5	Sequence 5, Appli
29	110.5	5.9	115	12	US-09-953-348-4	Sequence 4, Appli
30	110.5	5.9	115	15	US-10-267-255-4	Sequence 4, Appli
31	103.5	5.5	347	15	US-10-166-087-38	Sequence 38, Appl
32	98	5.2	555	10	US-09-801-368-350	Sequence 350, App
33	97.5	5.2	983	10	US-09-771-161A-227	Sequence 227, App
34	97.5	5.2	983	12	US-10-345-680-2	Sequence 2, Appli
35	97.5	5.2	983	15	US-10-205-823-97	Sequence 97, Appl
36	96.5	5.1	459	9	US-09-815-242-12703	Sequence 12703, A
37	96.5	5.1	952	15	US-10-216-556-2	Sequence 2, Appli
38	93.5	5.0	993	11	US-09-823-187-39	Sequence 39, Appl
39	93.5	5.0	993	11	US-09-823-187-41	Sequence 41, Appl
40	93.5	5.0	998	11	US-09-823-187-40	Sequence 40, Appl
41	93.5	5.0	998	11	US-09-823-187-42	Sequence 42, Appl
42	93.5	5.0	998	11	US-09-823-187-43	Sequence 43, Appl
43	93.5	5.0	1033	9	US-09-888-615-75	Sequence 75, Appl
44	92.5	4.9	841	9	US-09-861-451A-30	Sequence 30, Appl
45	90.5	4.8	821	12	US-10-205-219-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-10-361-460-5
; Sequence 5, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
; TITLE OF INVENTION: Biosynthesis and Uses Thereof
; FILE REFERENCE: 0709
; CURRENT APPLICATION NUMBER: US/10/361,460
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/057,082
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Zea mays
US-10-361-460-5

Query Match 31.3%; Score 590.5; DB 12; Length 375;
Best Local Similarity 37.3%; Pred. No. 7.3e-52;
Matches 139; Conservative 62; Mismatches 141; Indels 31; Gaps 12;

```
Qy          11 ASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTK 70
               : :: |    | | | |  |  |   ::| ||| || |  |   |  ::::   : |:|
Db           9 SQDLLQAHDELLHHSLCFAKSLALAVALDLRIPDAIHHHGAG-GATLLQILAETALHPSK 67

Qy          71 TRQVQSLMRYLAHNGFFEIVR----------IH----DNIEAYALTAASELLVKS--SE 113
               | :: ||| |   | | :|           :|    : :  | |||||  ||    |
Db          68 LRALRRLMRVLTVTGIFSVVEQPPAGGGDDSTVHTSDDEAVVVYRLTAASRFLVSDDVST 127

Qy         114 LSLAPMVEYFLEP--NCQGAWNQLKRWVHEEDLTVFEVSLG---TP-FWDFINKDPAYNK 167
               :||| |   |:|   |  |   :  |  :|      |    ||  |:  :    |
Db         128 ATLAPFVSLALQPIAACPHALG-ISAWFRQEQHEPSPYGLAFRQTPTIWEHADD---VNA 183
```

```
Qy        168 SFNEAMACDSQ-MLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERP 226
Db        184 LLNKGMAADSRFLMPIVLRECGETFRGIDSLVDVGGGHGGAAAAIAAAFPHLKCSVLDLP 243

Qy        227 NVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKT 286
Db        244 HVVAGAPSDGNVQFVAGNMFESIPPATAVFLKKTLHDWGDDECVKILKNCKQAISPRDAG 303

Qy        287 GKVVVVIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISP 346
Db        304 GKVIILDVVVGYKQSNIKHQETQVMFDLYMMA-VNGVERDEQEWKKIFTEAGFKDYKILP 362

Qy        347 FTGYLS-LIEIYP 358
Db        363 VIGDVSVIIEVYP 375
```

RESULT 2
US-10-361-460-6
; Sequence 6, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
; TITLE OF INVENTION: Biosynthesis and Uses Thereof
; FILE REFERENCE: 0709
; CURRENT APPLICATION NUMBER: US/10/361,460
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/057,082
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Zea mays
US-10-361-460-6

Query Match 30.1%; Score 567; DB 12; Length 370;
Best Local Similarity 34.7%; Pred. No. 1.8e-49;
Matches 137; Conservative 64; Mismatches 132; Indels 62; Gaps 11;

```
Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
Db          1 MALMQESSSQDQDMLQAHDELLHHSLCFAKSLALTVALDLRIPDAIHHHGGG--ATLLQI 58

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIH------DNIEA---YALTAASELLVK- 110
Db         59 LAETGLHPSKLRALRRLMRVLTVTGTFS-VQVQQPPAGSDDDEAVVVYRLTAASRFLVSD 117

Qy        111 --SSELSLAPMVEYFLEPNCQG-------AWNQLKR----------------WVHEEDLT 145
Db        118 EVSTATTLAPFVSLALQPIAASPHALGICAWFRQEQHEPSPYGLAFRQTPTLWEHADDV- 176

Qy        146 VFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQ-MLNLAFRDCNWVFEGLESIVDVGGGT 204
Db        177 -------------------NALLNKGMVADSRFLMPIVLRQCGEMFRGINSLVDVGGGH 216

Qy        205 GITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNW 264
Db        217 GGAAAAIAAAFPHVKCSVLDLPHVVAGAPSDGNVQFVAGNMFESIPPATAVFLKKTLHDW 276

Qy        265 NDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHMACIINGKE 324
Db        277 GDDECVKILKNCKQAIPPRDAGGKVIILDVVVGYKQSNIKHQETQVMFDLYMMA-VNGVE 335

Qy        325 RKEEDWKKLFMEAGFQSYKISPFTGYLS-LIEIYP 358
Db        336 RDEQEWKKIFAEAGFKDYKILPVIGDVSVIIEVYP 370
```

RESULT 3
US-10-361-460-7
; Sequence 7, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
; TITLE OF INVENTION: Biosynthesis and Uses Thereof
; FILE REFERENCE: 0709
; CURRENT APPLICATION NUMBER: US/10/361,460
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/057,082
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Zea mays
US-10-361-460-7

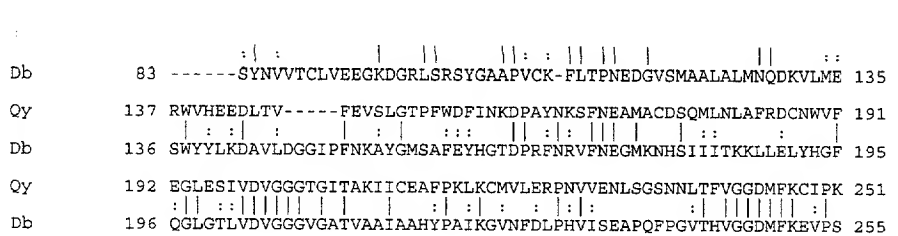
Query Match 28.7%; Score 540.5; DB 12; Length 366;
Best Local Similarity 33.3%; Pred. No. 9.5e-47;
Matches 128; Conservative 70; Mismatches 125; Indels 61; Gaps 11;

```
Qy         11 ASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTK 70
Db          8 SQDLLEAHDELFHHCLCFAKSLALAVAQDLRIPDAIHHHGGG--ATLHQILAEAALHPSK 65

Qy         71 TRQVQSLMRYLAHNGFFEIVRIHDNIEA-------YALTAASELLVKSSE----LSLAPM 119
Db         66 LRALRRLMRVLTVSGVF-TVQYSSTVDASDGADVVYRLTAASRFLVSDSDEAGTASLAPF 124

Qy        120 VEYFL-----EPNCQG--AWNQLKR----------------WVHEEDLTVFEVSLGTPFW 156
Db        125 ANLALHPIAISPHAVGICAWFRQEQHDPSPYGLAFRQIPTIWEHADNV------------ 172

Qy        157 DFINKDPAYNKSFNEAMACDSQ-MLNLAFRDC-NWVFEGLESIVDVGGGTGITAKIICEA 214
Db        173 ---------NALLNKGLLAESRFLMPIVLRECGDEVFRGIDSLVDVGGGHGGAAATIAAA 223

Qy        215 FPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILE 274
Db        224 FPHVKCSVLDLPHVVAGAPSDACVQFVAGNMFHSIPPATAVFFKTTLCDWGDDECIKILK 283

Qy        275 NCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLF 334
Db        284 NCKQAISPRDEGGKVIIMDVVVGYGQSNMKRLETQVMFDLVMMA-VNGVERDEQEWKEMF 342

Qy        335 MEAGFQSYKISPFTGYLSLIEIYP 358
Db        343 IEAGFKDYKIRPVAGLMSVIEVYP 366
```

RESULT 4
US-10-361-460-4
; Sequence 4, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
; TITLE OF INVENTION: Biosynthesis and Uses Thereof
; FILE REFERENCE: 0709
; CURRENT APPLICATION NUMBER: US/10/361,460
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/057,082
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 85

```
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 354
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-361-460-4

  Query Match             23.1%;  Score 435;  DB 12;  Length 354;
  Best Local Similarity   34.0%;  Pred. No. 6e-36;
  Matches  123;  Conservative   60;  Mismatches  147;  Indels   32;  Gaps   12;

Qy         12 SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKT 71
              :|: | || |::| | :: |  ||  ||| ||  |  |: |   | :||: | :| |
Db         10 AELLQAQADLWRHSLYYLTSMALKCAVELHIPTAI--HNLGGSATLPDLVAALSLPAAKL 67

Qy         72 RQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMVEYF----LEPN 127
               :  :|| |  :| |      |::: | |   | |||:  |       :||    :  :
Db         68 PFLGRVMRLLVTSGVF---ASSDDVQ-YRLNPLSWLLVEGVESEDHTYQKYFVLGTVSRH 123

Qy        128 CQGAWNQLKRWVHEED------LTVFEVSLGTPFWDFINK--DPAYNKSFNEAMAC-DSQ 178
              |   |  |  :|:       : ||   | |       |  |   ::   | :|  |:
Db        124 YVEAGMSLADWFKKEEDEDRQLPSPFEALHGVPLVHESTKLLDEELDRVVEEGVAAHDNL 183

Qy        179 MLNLAFRDCNW-VFEGLESIVDVGGGTG-ITAKIICEAFPKLKCMVLERPNVVENLSGSN 236
               :   |:|   || || |:    |  |  :|  | :||| :|| ||  | |||  : :
Db        184 AIGTVIRECGADVFSGLRSLTYCCGRQGNASAAAIVKAFPDIKCTVLNLPRVVEETT-TK 242

Qy        237 NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI 296
              :|       || | ||:|||||| |:|:||:|||| |::||    : |||::|: ::
Db        243 TIT---------IPPAQAVMLKLVLHFWSDDDCVKILELCRKAIPSRQEGGKVIIIEILL 293

Qy        297 NENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEI 356
                    : | :||||: |      |::| |:||: :| :|||  ||:    |  :||:
Db        294 GPYMGP-VMYEAQLLMDMLMMVNTKGRQRGEDDWRHIFTKAGFSDYKVVKKIGARGVIEV 352

Qy        357 YP 358
              ||
Db        353 YP 354


RESULT 5
US-09-947-027-6
; Sequence 6, Application US/09947027
; Patent No. US20020124281A1
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent Lee C.
;  APPLICANT:  Li, Laigeng
;  TITLE OF INVENTION: METHODS FOR SIMULTANEOUS CONTROL OF LIGNIN CONTENT AND COMPOSITIO
;  TITLE OF INVENTION:  CELLULOSE CONTENT IN PLANTS
;  FILE REFERENCE: 066040-9718
;  CURRENT APPLICATION NUMBER: US/09/947,027
;  CURRENT FILING DATE:  2001-09-05
;  PRIOR APPLICATION NUMBER: 60/230,086
;  PRIOR FILING DATE: 2000-09-05
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 365
;   TYPE: PRT
;   ORGANISM: aspen populus tremuloides
US-09-947-027-6

  Query Match             19.7%;  Score 371.5;  DB 10;  Length 365;
  Best Local Similarity   29.7%;  Pred. No. 2.1e-29;
  Matches   97;  Conservative   69;  Mismatches  132;  Indels   29;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFF-- 87
              ||  :|||: :|:      |  :: ||: | |   |||       :  ::| ||
Db         36 LKTAIELDLLEIMAKAGPGAFLSTSEIASHL---PTKNPDAPVMLDRILRLLASYSILTC 92

Qy         88 --------EIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRW 138
                     :: |:      | |    : | |: : :|::|:     :     :|  ||
Db         93 SLKDLPDGKVERL------YGLAPVCKFLTKNEDGVSVSPLCLMNQDKVLMESWYYLKDA 146

Qy        139 VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
              : :  :  |  : |   :::   || :|| ||: |:  | :       :    |||| |:|
Db        147 ILDGGIP-FNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYKGFEGLTSLV 205

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
              |||||||     |   :| :| :  : |:|:|:       :  ||||||  :|||||| :|
Db        206 DVGGGTGAVVNTIVSKYPSIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPKADAVFMK 265

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
              : |:|:|  |:| |:|| :|:    : |||:::: ::    |    |:  : :|| | |
Db        266 WICHDWSDAHCLKFLKNCYDAL---PENGKVILVECILPVAPDTSLATKGVVHVDVIMLA 322

Qy        318 CIINGKERKEEDWKKLFMEAGFQSYKI 344
                  |||| |:::: |   |||| :::
Db        323 HNPGGKERTEKEFEGLAKGAGFQGFEV 349


RESULT 6
US-10-091-009-6
; Sequence 6, Application US/10091009
; Publication No. US20020138870A1
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent Lee C.
;  APPLICANT:  Li, Laigeng
;  TITLE OF INVENTION: METHODS FOR SIMULTANEOUS CONTROL OF LIGNIN CONTENT AND
;  TITLE OF INVENTION:  COMPOSITION, AND
;  TITLE OF INVENTION:  CELLULOSE CONTENT IN PLANTS
;  FILE REFERENCE: 066040-9718
;  CURRENT APPLICATION NUMBER: US/10/091,009
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: 09/947,027
;  PRIOR FILING DATE: 2001-09-05
;  PRIOR APPLICATION NUMBER: 60/230,086
;  PRIOR FILING DATE: 2000-09-05
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 365
;   TYPE: PRT
;   ORGANISM: aspen populus tremuloides
US-10-091-009-6

  Query Match             19.7%;  Score 371.5;  DB 14;  Length 365;
  Best Local Similarity   29.7%;  Pred. No. 2.1e-29;
  Matches   97;  Conservative   69;  Mismatches  132;  Indels   29;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFF-- 87
              ||  :|||: :|:      |  :: ||: | |   |||       :  ::| ||
Db         36 LKTAIELDLLEIMAKAGPGAFLSTSEIASHL---PTKNPDAPVMLDRILRLLASYSILTC 92

Qy         88 --------EIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRW 138
                     :: |:      | |    : | |: : :|::|:     :     :|  ||
Db         93 SLKDLPDGKVERL------YGLAPVCKFLTKNEDGVSVSPLCLMNQDKVLMESWYYLKDA 146

Qy        139 VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
              : :  :  |  : |   :::   || :|| ||: |:  | :       :    |||| |:|
Db        147 ILDGGIP-FNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYKGFEGLTSLV 205

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
              |||||||     |   :| :| :  : |:|:|:       :  ||||||  :|||||| :|
Db        206 DVGGGTGAVVNTIVSKYPSIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPKADAVFMK 265

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
              : |:|:|  |:| |:|| :|:    : |||:::: ::    |    |:  : :|| | |
Db        266 WICHDWSDAHCLKFLKNCYDAL---PENGKVILVECILPVAPDTSLATKGVVHVDVIMLA 322

Qy        318 CIINGKERKEEDWKKLFMEAGFQSYKI 344
                  |||| |:::: |   |||| :::
Db        323 HNPGGKERTEKEFEGLAKGAGFQGFEV 349


RESULT 7
```

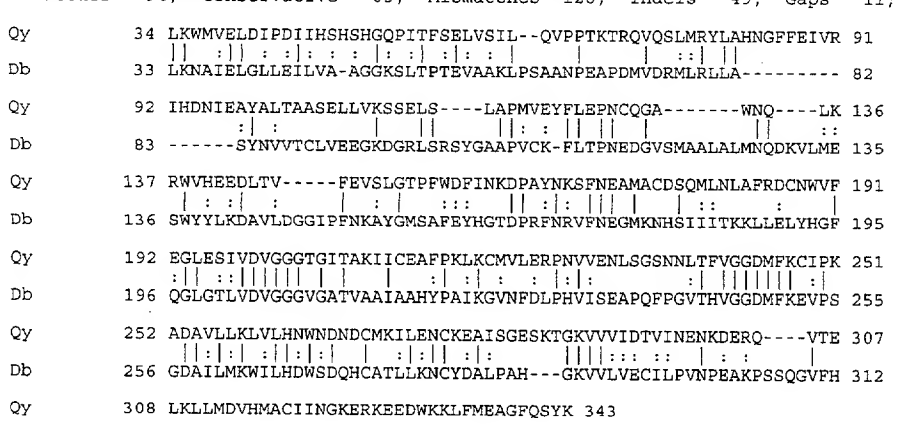
```
US-09-796-256A-6
; Sequence 6, Application US/09796256A
; Patent No. US20020078477A1
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent L
;  APPLICANT:  Carraway, Daniel T
;  APPLICANT:  Smeltzer, Richard H
;  TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
;  FILE REFERENCE: 50617/c-3532.0
;  CURRENT APPLICATION NUMBER: US/09/796,256A
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER: US60/033381
;  PRIOR FILING DATE: 1996-12-16
;  PRIOR APPLICATION NUMBER: 08/991677
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 368
;   TYPE: PRT
;   ORGANISM: Liquidambar styraciflua
US-09-796-256A-6

  Query Match             18.6%;  Score 350.5;  DB 9;  Length 368;
  Best Local Similarity   29.0%;  Pred. No. 3e-27;
  Matches  100;  Conservative   62;  Mismatches  130;  Indels   53;  Gaps   11;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFFEI 89
              ||  :|||: :|:      |  |: |:: | |   |||       :  ::| ||
Db         38 LKSAIELDVLEIMAKAGPGAHISTSDIASKL---PTKNPDAAVMLDRMLRLLA------- 87

Qy         90 VRIHDNIEAYALTAASELLVKSSEL----SLAPMVEYFLEPNCQGA-------WNQ---- 134
                 :|::   |   :  ::     |||: : ||  |  |                ||
Db         88 --------SYSVLTCSLRTLPDGKIERLYGLAPVCK-FLTRNDDGVSIAALSLMNQDKVL 138

Qy        135 LKRWVHEEDLTVFEVSLGTPF--------WDFINKDPAYNKSFNEAMACDSQMLNLAFRD 186
              :: | |   ||   :  | ||        :::   || :|  ||  |:  | :      :
Db        139 MESWYH---LTEAVLEGGIPFNKAYGMTAFEYHGTDPRFNTVFNNGMSNHSTITMKKILE 195

Qy        187 CNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMF 246
                |||| |:|||||||    :|   :| :| :  :  |:|:|       :  ||||||
Db        196 TYKGFEGLGSVVDVGGGTGAHLNMIIAKYPMIKGINFDLPHVIEEAPSYPGVEHVGGDMF 255

Qy        247 KCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVT 306
               :|| ||: :| : |:|:|   |:| |: | ||:       |||:: : ::    |  |
Db        256 VSVPKGDAIFMKWICHDWSDEHCLKFLKKCYEAL---PTNGKVILAECILPVAPDASLPT 312

Qy        307 ELKLLMDVHM-ACIINGKERKEEDWKKLFMEAGFQSYKISPFTGY 350
              :  : :|| | |    |||| |:::: |   |||: :::     |
Db        313 KAVVHIDVIMLAHNPGGKERTEKEFEALAKGAGFEGPRVVASCAY 357


RESULT 8
US-10-289-757-103
; Sequence 103, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
;  APPLICANT: Demmer, Jeroen
;  APPLICANT:  Forster, Richard L
;  APPLICANT:  Gibson, John Bryan
;  APPLICANT:  Shenk, Michael Andrew
;  APPLICANT:  No. US20030180751A1riss, Geoffrey
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Saulsbury, Keith Martin
;  APPLICANT:  Hall, Claire
;  TITLE OF INVENTION: Compositions isolated from forage
;  TITLE OF INVENTION:  grasses and methods for their use
;  FILE REFERENCE: 11000.1061U
;  CURRENT APPLICATION NUMBER: US/10/289,757
;  CURRENT FILING DATE:  2002-11-07
;  PRIOR APPLICATION NUMBER: 60/337,703
;  PRIOR FILING DATE: 2001-11-07
;  NUMBER OF SEQ ID NOS: 218
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103
;   LENGTH: 360
;   TYPE: PRT
;   ORGANISM: FLolium perenne
US-10-289-757-103

  Query Match             17.9%;  Score 337.5;  DB 12;  Length 360;
  Best Local Similarity   28.6%;  Pred. No. 6.3e-26;
  Matches   96;  Conservative   62;  Mismatches  129;  Indels   49;  Gaps   11;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL--QVPPTKTRQVQSLMRYLAHNGFFEIVR 91
              ||  :|| : :|: : : |: :| :|: : |      |      |  ::| ||
Db         33 LKNAIELGLLEILVA-AGGKSLTPTEVAAKLPSAANPEAPDMVDRILRLLA--------- 82

Qy         92 IHDNIEAYALTAASELLVKSSELS----LAPMVEYFLEPNCQGA-------WNQ----LK 136
                    :| :      |   ||      ||: : || ||  |          ||    ::
Db         83 ------SYNVVTCLVEEGKDGRLSRSYGAAPVCK-FLTPNEDGVSMAALALMNQDKVLME 135

Qy        137 RWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVF 191
              | : :|  :      |  : |   :::   || :|: ||| |   | ::     :   |
Db        136 SWYYLKDAVLDGGIPFNKAYGMSAFEYHGTDPRFNRVFNEGMKNHSIIITKKLLELYHGF 195

Qy        192 EGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK 251
              ||| |:|||||| | |   |  :| :| :  : |:|:        :| |||||| :|
Db        196 EGLGSLVDVGGGVGATVAAIAAHYPTIKGVNFDLPHVISEAPQFPGVTHVGGDMFKEVPS 255

Qy        252 ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI----NENKDERQVTE 307
              ||:|:| :||:|:|  |  :|:|| :|:      ||||::: ::      |   : |
Db        256 GDAILMKWILHDWSDQHCATLLKNCYDALPAH---GKVVLVECILPVNPEANPSSQGVFH 312

Qy        308 LKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYK 343
              : ::|  |      |:|| | ::: |   |||   |
Db        313 VDMIMLAHNP---GGRERYEREFQALARGAGFTGVK 345


RESULT 9
US-10-289-757-102
; Sequence 102, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
;  APPLICANT: Demmer, Jeroen
;  APPLICANT:  Forster, Richard L
;  APPLICANT:  Gibson, John Bryan
;  APPLICANT:  Shenk, Michael Andrew
;  APPLICANT:  No. US20030180751A1riss, Geoffrey
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Saulsbury, Keith Martin
;  APPLICANT:  Hall, Claire
;  TITLE OF INVENTION: Compositions isolated from forage
;  TITLE OF INVENTION:  grasses and methods for their use
;  FILE REFERENCE: 11000.1061U
;  CURRENT APPLICATION NUMBER: US/10/289,757
;  CURRENT FILING DATE:  2002-11-07
;  PRIOR APPLICATION NUMBER: 60/337,703
;  PRIOR FILING DATE: 2001-11-07
;  NUMBER OF SEQ ID NOS: 218
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 102
;   LENGTH: 360
;   TYPE: PRT
;   ORGANISM: Festuca arundinacea
US-10-289-757-102

  Query Match             17.6%;  Score 331.5;  DB 12;  Length 360;
  Best Local Similarity   28.0%;  Pred. No. 2.6e-25;
  Matches   94;  Conservative   65;  Mismatches  128;  Indels   49;  Gaps   11;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL--QVPPTKTRQVQSLMRYLAHNGFFEIVR 91
              ||  :|| : :|: : : |: :| :|: : |      |      |  ::| ||
Db         33 LKNAIELGLLEILVA-AGGKSLTPTEVAAKLPSAANPEAPDMVDRMLRLLA--------- 82

Qy         92 IHDNIEAYALTAASELLVKSSELS----LAPMVEYFLEPNCQGA-------WNQ----LK 136
```



```
Db          83 ------SYNVVTCLVEEGKDGRLSRSYGAAPVCK-FLTPNEDGVSMAALALMNQDKVLME 135

Qy         137 RWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVF 191

Db         136 SWYYLKDAVLDGGIPFNKAYGMSAFEYHGTDPRFNRVFNEGMKNHSIIITKKLLELYHGF 195

Qy         192 EGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK 251

Db         196 QGLGTLVDVGGGVGATVAAIAAHYPAIKGVNFDLPHVISEAPQFPGVTHVGGDMFKEVPS 255

Qy         252 ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQ----VTE 307

Db         256 GDAILMKWILHDWSDQHCATLLKNCYDALPAH---GKVVLVECILPVNPEAKPSSQGVFH 312

Qy         308 LKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYK 343

Db         313 VDMIMLAHNP---GGRERYEREFEALARGAGFTGVK 345
```

RESULT 10
US-10-289-757-182
; Sequence 182, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forster, Richard L
; APPLICANT: Gibson, John Bryan
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: No. US20030180751A1riss, Geoffrey
; APPLICANT: Glenn, Matthew
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Compositions isolated from forage
; TITLE OF INVENTION: grasses and methods for their use
; FILE REFERENCE: 11000.1061U
; CURRENT APPLICATION NUMBER: US/10/289,757
; CURRENT FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/337,703
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 182
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Festuca arundinacea
US-10-289-757-182

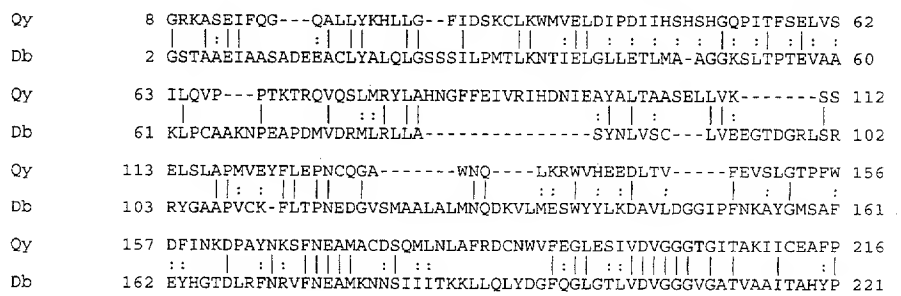
Query Match 17.6%; Score 331.5; DB 12; Length 360;
Best Local Similarity 28.0%; Pred. No. 2.6e-25;
Matches 94; Conservative 65; Mismatches 128; Indels 49; Gaps 11;



```
Qy          34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL--QVPPTKTRQVQSLMRYLAHNGFFEIVR 91

Db          33 LKNAIELGLLEILVA-AGGKSLTPTEVAAKLPSAANPEAPDMVDRMLRLLA--------- 82

Qy          92 IHDNIEAYALTAASELLVKSSELS----LAPMVEYFLEPNCQGA-------WNQ----LK 136

Db          83 ------SYNVVTCLVEEGKDGRLSRSYGAAPVCK-FLTPNEDGVSMAALALMNQDKVLME 135

Qy         137 RWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVF 191

Db         136 SWYYLKDAVLDGGIPFNKAYGMSAFEYHGTDPRFNRVFNEGMKNHSIIITKKLLELYHGF 195

Qy         192 EGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK 251

Db         196 QGLGTLVDVGGGVGATVAAIAAHYPAIKGVNFDLPHVISEAPQFPGVTHVGGDMFKEVPS 255

Qy         252 ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQ----VTE 307

Db         256 GDAILMKWILHDWSDQHCATLLKNCYDALPAH---GKVVLVECILPVNPEAKPSSQGVFH 312

Qy         308 LKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYK 343

Db         313 VDMIMLAHNP---GGRERYEREFEALARGAGFTGVK 345
```

RESULT 11
US-10-289-757-105
; Sequence 105, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forster, Richard L
; APPLICANT: Gibson, John Bryan
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: No. US20030180751A1riss, Geoffrey
; APPLICANT: Glenn, Matthew
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Compositions isolated from forage
; TITLE OF INVENTION: grasses and methods for their use
; FILE REFERENCE: 11000.1061U
; CURRENT APPLICATION NUMBER: US/10/289,757
; CURRENT FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/337,703
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Lolium perenne
US-10-289-757-105

Query Match 17.5%; Score 329.5; DB 12; Length 361;
Best Local Similarity 27.8%; Pred. No. 4.2e-25;
Matches 103; Conservative 70; Mismatches 137; Indels 61; Gaps 14;



```
Qy           8 GRKASEIFQG---QALLYKHLLG--FIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVS 62

Db           2 GSTAAEIAASADEEACLYALQLGSSSILPMTLKNTIELGLLETLMA-AGGKSLTPTEVAA 60

Qy          63 ILQVP---PTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVK-------SS 112

Db          61 KLPCAAKNPEAPDMVDRMLRLLA--------------SYNLVSC---LVEEGTDGRLSR 102

Qy         113 ELSLAPMVEYFLEPNCQGA-------WNQ----LKRWVHEEDLTV-----FEVSLGTPFW 156

Db         103 RYGAAPVCK-FLTPNEDGVSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMSAF 161

Qy         157 DFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFP 216

Db         162 EYHGTDLRFNRVFNEAMKNNSIIITKKLLQLYDGFQGLGTLVDVGGGVGATVAAITAHYP 221

Qy         217 KLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENC 276

Db         222 TIKGINFDLPHVISEAPPFPGVTHIGGDMFKKVPSGDAILMKWILHDWSDQHCATLLKNC 281

Qy         277 KEAISGESKTGKVVVIDTVINENKDERQ----VTELKLLMDVHMACIINGKERKEEDWKK 332

Db         282 YDALPVH---GKVVLVECILPVNPEAKPSSQGVFHVDMIMLAHNP---GGRERYEREYEA 335

Qy         333 LFMEAGFQSYK 343

Db         336 LARGAGFAGFK 346
```

RESULT 12
US-10-289-757-104
; Sequence 104, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forster, Richard L
; APPLICANT: Gibson, John Bryan

```
;  APPLICANT:  Shenk, Michael Andrew
;  APPLICANT:  No. US20030180751A1riss, Geoffrey
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Saulsbury, Keith Martin
;  APPLICANT:  Hall, Claire
;  TITLE OF INVENTION: Compositions isolated from forage
;  TITLE OF INVENTION:  grasses and methods for their use
;  FILE REFERENCE: 11000.1061U
;  CURRENT APPLICATION NUMBER: US/10/289,757
;  CURRENT FILING DATE:  2002-11-07
;  PRIOR APPLICATION NUMBER: 60/337,703
;  PRIOR FILING DATE: 2001-11-07
;  NUMBER OF SEQ ID NOS: 218
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104
;   LENGTH: 360
;   TYPE: PRT
;   ORGANISM: Festuca arundinacea
US-10-289-757-104

  Query Match             17.4%;  Score 328.5;  DB 12;  Length 360;
  Best Local Similarity   27.7%;  Pred. No. 5.2e-25;
  Matches   93;  Conservative   67;  Mismatches  127;  Indels   49;  Gaps   11;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL--QVPPTKTRQVQSLMRYLAHNGFFEIVR 91
Db         33 LKNAIELGLLEILVA-AGGKSLTPTEVAAKLPSAANPEAPDMVDRMLRLLA--------- 82

Qy         92 IHDNIEAYALTAASELLVKSSELS----LAPMVEYFLEPNCQGA-------WNQ----LK 136
Db         83 ------SYNVVSCLVEEGKDGRLSRNYGAAPVCK-FLTPNEDGVSMAALALMNQDKVLME 135

Qy        137 RWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVF 191
Db        136 SWYYLKDAVLDGGIPFNKAYGMSAFEYHGTDPRFNRVFNEGMKNHSIIITKKLLELYDGF 195

Qy        192 EGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK 251
Db        196 QGLGTLVDVGGGVGATVAAITAHYPAIKGVNFDLPHVISEAPPFPGVTHVGGDMFKKVPS 255

Qy        252 ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQ----VTE 307
Db        256 GDAIMMKWILHDWSDQHCATLLKNCYDALPAH---GKVVLVECILPVNPEAKPSSQGVFH 312

Qy        308 LKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYK 343
Db        313 VDMIMLAHNP---GGRERYEREFEALARGAGFAGVK 345

RESULT 13
US-10-213-473-29
; Sequence 29, Application US/10213473
; Publication No. US20030104598A1
; GENERAL INFORMATION:
;  APPLICANT: Rathinasabapathi, Bala
;  APPLICANT:  Raman, Suresh Babu
;  TITLE OF INVENTION: BETA-ALANINE N-METHYLTRANSFERASE
;  FILE REFERENCE: 5853-205-1CIP
;  CURRENT APPLICATION NUMBER: US/10/213,473
;  CURRENT FILING DATE:  2002-08-06
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
;   LENGTH: 375
;   TYPE: PRT
;   ORGANISM: Limonium latifolium
US-10-213-473-29

  Query Match             15.5%;  Score 293;  DB 15;  Length 375;
  Best Local Similarity   25.6%;  Pred. No. 2.4e-21;
  Matches   87;  Conservative   69;  Mismatches  154;  Indels   30;  Gaps    9;

Qy         37 MVELDIPDIIHSHSHGQPITFSELVSILQVPP--TKTRQVQSLMRYLAHNGFFEIVRIHD 94
Db         42 IVRLKVPEAIWSNGSNTPVSAAEILSRLPDAPATADAENLQRLLRVLTSFGVFS-----E 96

Qy         95 NIEA------------YALTAASELLVKSSELSLAPMVEYFLEPNCQ---GAWNQLKRWV 139
Db         97 HLDTTSSSSSSTSERRYCLTEVGQTLVSFDE-SCPSHGAYVLQHHQETLLKAWPFLHTAI 155

Qy        140 HEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNW-VFEGLESIV 198
Db        156 LDASTEPFARVNGEPAYQYYGKNDELNKNMQYAMSGVSVPYMKALLGSGYDGFEGVKTLV 215

Qy        199 DVGGGTGITAKIICEAFPKL-KCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLL 257
Db        216 DVGGSSGDCLRMIINKYKDIPKAINFDLPEVVAKAPKIPGITHVGGNMFESVPSGDAIFV 275

Qy        258 KLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHMA 317
Db        276 KWVLTCFTDEEVITLMRNCNKAL---PVGGKLICSEPTLPENSDESHRTRALLVADIFIM 332

Qy        318 CI--INGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIE 355
Db        333 TTYRAKGKHRTEEEYRQLGLLAGFPKFRVIHVDYFFPVVE 372

RESULT 14
US-10-174-693-274
; Sequence 274, Application US/10174693
; Publication No. US20030131373A1
; GENERAL INFORMATION:
;  APPLICANT: Bloksberg, Leonard N.
;  APPLICANT:  Havukkala, Ilkka
;  TITLE OF INVENTION: Materials and Methods for the
;  TITLE OF INVENTION:  Modification of Plant Lignin Content
;  FILE REFERENCE: 11000.1003c5
;  CURRENT APPLICATION NUMBER: US/10/174,693
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 08/975,316
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: US 08/713,000
;  PRIOR FILING DATE: 1996-09-11
;  PRIOR APPLICATION NUMBER: US 09/169,789
;  PRIOR FILING DATE: 1998-10-09
;  PRIOR APPLICATION NUMBER: US 09/615,192
;  PRIOR FILING DATE: 2000-07-12
;  NUMBER OF SEQ ID NOS: 407
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 274
;   LENGTH: 328
;   TYPE: PRT
;   ORGANISM: Eucalyptus grandis
US-10-174-693-274

  Query Match             15.0%;  Score 283.5;  DB 16;  Length 328;
  Best Local Similarity   28.0%;  Pred. No. 1.9e-20;
  Matches   79;  Conservative   50;  Mismatches  142;  Indels   11;  Gaps    4;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVP-PTKTRQVQSLMRYLAHNGFFEIVRI 92
Db         53 LKGVHELGILELL---AKGDQLSPLDIVARLSIDNPAAPDTIDRMLRLLASYSILSCTLV 109

Qy         93 HDN----IEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQLKRWVHEEDLTVFE 148
Db        110 EDKEGRPQRLYGLGPRSKFFLDQNGASTLPTHMLLQEKTLLECWNCLKDAVKEGGADPFT 169

Qy        149 VSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGGTGITA 208
Db        170 RRHGMNVFDYMGQDPRFNDLYNKSMRTGSAIYMPKIAQHYRGFSKAKTVVNVGGGIGETL 229

Qy        209 KIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDND 268
Db        230 KTILSKNPHIRAINYDLPHVIATAPPIPGITHVGGDILKSVPKADVHFLKSVLHRGDDEF 289

Qy        269 CMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKL 310
```

```
                |:|:|:|| ||:     |||||::: |  |      |::  |
Db        290 CVKVLKNCWEAL---PPTGKVVIVEEVTPEYPGTDDVSQTTL 328


RESULT 15
US-09-953-348-133
; Sequence 133, Application US/09953348
; Publication No. US20030134398A1
; GENERAL INFORMATION:
;  APPLICANT: Sherman, David. H
;  APPLICANT:  Mao, Yingqing
;  APPLICANT:  Varoglu, Mustafa
;  APPLICANT:  He, Min
;  APPLICANT:  Sheldon, Paul
;  TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
;  FILE REFERENCE: 600.530US1
;  CURRENT APPLICATION NUMBER: US/09/953,348
;  CURRENT FILING DATE:  2001-09-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06394
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/266965
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 133
;   LENGTH: 351
;   TYPE: PRT
;   ORGANISM: Streptomyces lavendulae
US-09-953-348-133

  Query Match             14.5%;  Score 273.5;  DB 12;  Length 351;
  Best Local Similarity   27.0%;  Pred. No. 2.2e-19;
  Matches   95;  Conservative   68;  Mismatches  140;  Indels   49;  Gaps   15;

Qy         25 LLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSLMRYLAHN 84
              | |    :: :   ||| :|:::       | | : |    :      : :: |:| ||
Db         26 LQGAWKARAIHVAVELGVPELLQE----GPRTATALA---EATGAHEQTLRRLLRLLATV 78

Qy         85 GFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQLKRWVHEEDL 144
              | |: :  ||::  :|  | | :|:      :|    :   |    || ||   |   :
Db         79 GVFDDLG-HDDL--FAQNALSAVLLPDPASPVATDARFQAAPWHWRAWEQLTHSVRTGEA 135

Qy        145 TVFEV------SLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWV-----FEG 193
              :          |  :|      :||   : || ||      :::  :   |     | |
Db        136 SFPSTWPTAPRSGSSP-----TRDPKARELFNRAMG------SVSLTEAGQVAAAYDFSG 184

Qy        194 LESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVEN----LSG---SNNLTFVGGDMF 246
              : ||:||| |     : :||| |:  :|||| | |     |:|   ::    : || |
Db        185 AATAVDIGGGRGSLMAAVLDAFPGLRGTLLERPPVAEEARELLTGRGLADRCEILPGDFF 244

Qy        247 KCIPK-ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQV 305
              : ||  ||  |:| |||:|:|:| ::||     |:  :|   :::||| :|   |||
Db        245 ETIPDGADVYLIKHVLHDWDDDDVVRILRRIATAMKPDS---RLLVIDNLI----DERPA 297

Qy        306 TELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPF-TGYLSLIEI 356
                | :|: :  :: | || | ::  |  ::| :  :  |   | : ::||
Db        298 AS-TLFVDLLLLVLVGGAERSESEFAALLEKSGLRVERSLPCGAGPVRIVEI 348


Search completed: November 10, 2003, 03:30:38
Job time : 77 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 10, 2003, 03:09:40 ; Search time 43 Seconds
(without alignments)
800.660 Million cell updates/sec

Title: US-09-868-547-4
Perfect score: 1886
Sequence: 1 MASSLNNGRKASEIFQGQAL.....FQSYKISPFTGYLSLIEIYP 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1295	68.7	352	2	T09254	isoflavone-7-O-met
2	1285	68.1	352	2	T09707	isoflavone-O-methy
3	1237.5	65.6	343	2	T09299	o-methyltransferas
4	867.5	46.0	360	2	T06786	6a-hydroxymaackiai
5	652.5	34.6	390	2	S52015	catechol O-methylt
6	606	32.1	364	2	JQ2268	O-methyltransferas
7	529.5	28.1	382	2	T09600	catechol O-methylt
8	527	27.9	382	2	T04963	catechol O-methylt
9	485	25.7	325	2	T04962	catechol O-methylt
10	426	22.6	363	2	T01354	herbicide safener
11	390.5	20.7	364	2	S36403	catechol O-methylt
12	386.5	20.5	364	2	S36404	catechol O-methylt
13	384	20.4	372	2	T09617	isoliquiritigenin
14	383.5	20.3	359	2	T12259	O-diphenol-O-methy
15	379.5	20.1	365	2	T09673	caffeate O-methylt
16	371.5	19.7	365	2	S18568	lignin-bispecific
17	362.5	19.2	364	2	T09780	probable caffeate
18	358.5	19.0	366	2	S40146	catechol O-methylt
19	351	18.6	365	2	JQ2344	catechol O-methylt
20	342	18.1	363	2	E96559	hypothetical prote
21	327.5	17.4	364	2	S28612	catechol O-methylt
22	326	17.3	376	2	T06189	probable catechol
23	322.5	17.1	350	2	T12260	caffeoyl-CoA O-met
24	318.5	16.9	367	2	E96796	hypothetical prote
25	316.5	16.8	352	2	H86454	CDS protein F9L11.
26	303	16.1	359	2	T46160	caffeic acid O-met
27	303	16.1	365	2	S22696	myo-inositol O-met
28	299.5	15.9	381	2	H96656	hypothetical prote
29	283.5	15.0	381	2	F96804	hypothetical prote
30	283.5	15.0	381	2	G96804	hypothetical prote
31	278.5	14.8	373	2	B86344	hypothetical prote
32	278.5	14.8	373	2	E86344	hypothetical prote
33	272.5	14.4	373	2	C86344	hypothetical prote
34	271.5	14.4	373	2	D86344	probable O-methylt
35	268	14.2	356	2	A47128	carminomycin 4-O-m
36	257	13.6	334	2	F83120	probable O-methylt
37	252	13.4	376	2	JQ1393	O-demethylpuromyci
38	245.5	13.0	205	2	E96653	hypothetical prote
39	240.5	12.8	341	2	JC5855	polyketide synthas
40	202	10.7	494	2	S27696	tcmN protein - Str
41	200.5	10.6	339	2	F70932	hypothetical prote
42	194	10.3	345	2	A42106	acetylserotonin O-
43	192	10.2	346	2	S21265	acetylserotonin O-
44	176	9.3	373	2	I37463	acetylserotonin O-
45	170.5	9.0	379	2	T50751	hydroxyneurosporen

ALIGNMENTS

RESULT 1
T09254
isoflavone-7-O-methyltransferase (EC 2.1.1.-) 9 - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: T09254
R;He, X.Z.; Reddy, J.T.; Dixon, R.A.
Plant Mol. Biol. 36, 43-54, 1998
A;Title: Stress responses in alfalfa (Medicago sativa L.) XXII. cDNA cloning and chara
A;Reference number: Z16628; MUID:98145455; PMID:9484461
A;Accession: T09254
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-352 <HEX>
A;Cross-references: EMBL:AF000976; NID:g2580583; PIDN:AAC49927.1; PID:g2580584
C;Genetics:
A;Note: 7-IOMT(9)
C;Function:
A;Description: catalyzes the O-methylation of A-ring hydroxyl group(s) of isoflavones;
A;Note: elicitor-inducible
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

Query Match 68.7%; Score 1295; DB 2; Length 352;
Best Local Similarity 68.2%; Pred. No. 2.7e-96;
Matches 244; Conservative 55; Mismatches 53; Indels 6; Gaps 4;

```
Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGPIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              |||| |||| |||| |||||||  |||| ||| | ::||:|| |:||:||: | |
Db          1 MASSI-NGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVGMNIPNII--HNHGKPISLSNL 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120
              ||||||| :|   |: |||||||||||||:   :  |:|||| |||||| |:| |||||
Db         58 VSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEE--ESYALTVASELLVRGSDLCLAPMV 115

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDPINKDPAYNKSFNEAMACDSQML 180
              |  |:|   |::::||:|::|||||:| |:||: ||||::|:| || |||:||| ||:::
Db        116 ECVLDPTLSGGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMASDSKLI 175

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240
              ||| |||::||:|||||||||||||| |||||| ||||||:| :|| ||||||||||||:
Db        176 NLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTY 235

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300
              ||||||  || ||||||| :|||| | ||::||: ||||:: : | ||| :|| |||| |
Db        236 VGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVTNDGKRGKVTIIDMVINEKK 295

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              || |||::|||||||:||| :||||| ||:|||||:||||| ||||| ||:|||||||
Db        296 DENQVTQIKLLMDVNMAC-LNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLIEIYP 352
```

RESULT 2
T09707
isoflavone-O-methytransferase (EC 2.1.1.-) - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Nov-2000
C;Accession: T09707; T09250
R;He, X.Z.; Reddy, J.T.; Dixon, R.A.
Plant Mol. Biol. 36, 43-54, 1998
A;Title: Stress responses in alfalfa (Medicago sativa L.) XXII. cDNA cloning and charact
A;Reference number: Z16628; MUID:98145455; PMID:9484461
A;Accession: T09707
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-352 <HEX>
A;Cross-references: EMBL:U97125; NID:g2580539; PIDN:AAC49928.1; PID:g2580540
A;Accession: T09250
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-292,'E',294-352 <HE2>
A;Cross-references: EMBL:AF000975; NID:g2580581; PIDN:AAC49926.1; PID:g2580582
C;Genetics:
A;Note: 7-IOMT(6)
C;Function:
A;Description: methyltransferase
A;Note: elicitor-inducible
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

Query Match 68.1%; Score 1285; DB 2; Length 352;
Best Local Similarity 67.6%; Pred. No. 1.7e-95;
Matches 242; Conservative 57; Mismatches 53; Indels 6; Gaps 4;

```
Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              ||||: |||| ||||: ||||||:  |||| ||| ||::||:||   :||:||: | |
Db          1 MASSI-NGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNII--QNHGKPISLSNL 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120
              |||||| :|   |: |||||||||||||:   :  |:|||| ||||||: |:| |||||
Db         58 VSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEE--ESYALTVASELLVRGSDLCLAPMV 115

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180
              | |:|   |::::||:|::|||||:| |:|: ||||::|:| || |||:||| ||:::
Db        116 ECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMASDSKLI 175

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240
              ||| |||::||:|||||||||||| ||||||| ||||||:| :|| ||||||||||||:
Db        176 NLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTY 235

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300
              ||||||  || ||||||| :|||| | ||::||: |||||:: : | ||| :|| ||:: |
Db        236 VGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVTNDGKRGKVTIIDMVIDKKK 295

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              || |||::||||||:||| :||||| ||:|||||:||||| ||||| ||:|||||||
Db        296 DENQVTQIKLLMDVNMAC-LNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLIEIYP 352
```

RESULT 3
T09299
o-methyltransferase (EC 2.1.1.-) iomt2003 - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 26-Aug-1999
C;Accession: T09299
R;He, X.Z.; Dixon, R.A.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z16642
A;Accession: T09299
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-343 <HEX>
A;Cross-references: EMBL:AF023481; NID:g2565272; PID:g2565273
C;Genetics:
A;Note: iomt2003
C;Function:
A;Description: methyltransferase
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

Query Match 65.6%; Score 1237.5; DB 2; Length 343;
Best Local Similarity 66.5%; Pred. No. 1.1e-91;
Matches 238; Conservative 55; Mismatches 50; Indels 15; Gaps 5;

```
Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              ||||: |||| ||||: ||||||:  |||| ||| ||::||:||   :||:||: | |
Db          1 MASSI-NGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNII--QNHGKPISLSNL 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120
              ||||||| :|   |: |||||||||||||:   :  |:|||| ||||||: |:| |||||
Db         58 VSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEE--ESYALTVASELLVRGSDLCLAPMV 115

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180
              | |:|   |::::||:|::|||||:| |:|: ||||::|:| || |||:||| ||:::
Db        116 ECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMASDSKLI 175

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240
              ||| |||::||:|||||||||||| ||||||| ||||||:| :|| ||||||||||||:
Db        176 NLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTY 235

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300
              ||||||  || ||||||| :|||| | ||::||: |||||:: : | ||| :|| ||:: |
Db        236 VGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVTNDGKRGKVTIIDMVIDKKK 295

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              || |||::||||||:||| :||||| ||:|||||:||||| ||||| ||:|||||||||
Db        296 DENQVTQIKLLMDVNMAC-LNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLIEIYP 352
```

RESULT 4
T06786
6a-hydroxymaackiain methyltransferase (EC 2.1.1.-) - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: T06786
R;Wu, Q.; Preisig, C.L.; VanEtten, H.D.
Plant Mol. Biol. 35, 551-560, 1997
A;Title: Isolation of the cDNAs encoding (+)6a-hydroxymaackiain 3-O-methyltransferase,
A;Reference number: Z15813; MUID:98009990; PMID:9349277
A;Accession: T06786
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-360 <WUQ>
A;Cross-references: EMBL:U69554; NID:g1568636; PIDN:AAC49856.1; PID:g1568637
A;Experimental source: cv. Alaska
C;Genetics:
A;Note: hmm6
C;Function:
A;Description: catalyzes the methylation of (+)6a-hydroxymaackiain to (+)pisatin
A;Pathway: the terminal step in the pisatin biosynthesis
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

Query Match 46.0%; Score 867.5; DB 2; Length 360;
Best Local Similarity 48.7%; Pred. No. 5.8e-62;
Matches 174; Conservative 60; Mismatches 116; Indels 7; Gaps 3;

```
Qy          7 NGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQV 66
              || : ||::  |  ||||:  |: |  ||  :|| | | |  |:||:|:|  || | |::
Db          6 NGSEESELYHAQIHLYKHVYNFVSSMALKSAMELGIADAI--HNHGKPMTLPELSSSLKL 63

Qy         67 PPTKTRQVQSLMRYLAHNGFFEIVRIHDN----IEAYALTAASELLVKSSELSLAPMVEY 122
              |:|   :  :| | |||||   :  |      || || :|:|||      |: :|:
Db         64 HPSKVNILYRFLRLLTHNGFFAKTTVKSNEGEEETAYVLTPSSKLLVSGKSTCLSSLVKG 123

Qy        123 FLEPNCQGAWNQLKRWVHEE-DLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLN 181
              | |:    |  |:| ||: : |:|| : |  :|||:|||      | :||| ||::
Db        124 ALHPSSLDMWGVSKKWFHEDKEQTLFECATGENYWDFLNKDSDSLSMFQDAMAADSRLFK 183
```

```
Qy        182 LAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFV 241
              ||  ::   |||||||:||| ||||  ||:| |||| :|| | ::| || ||:|: || ||
Db        184 LAIQENKHVFEGLESLVDVAGGTGGVAKLIHEAFPHIKCTVFDQPQVVGNLTGNENLNFV 243

Qy        242 GGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD 301
              |||||| :| ||||||| |||:||| :|||:| ||||| : | |||::||  |:|| |
Db        244 GGDMFKSVPSADAVLLKWVLHDWNDELSLKILKNSKEAISHKGKDGKVIIIDISIDENSD 303

Qy        302 ERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :| :|||:|  || |  :  |||| :::|:|| :||| |||:|  |: ||||:||
Db        304 DRGLTELQLEYDVVMLTMFLGKERTKKEWEKLIYDAGFSRYKITPICGFKSLIEVYP 360


RESULT 5
S52015
catechol O-methyltransferase (EC 2.1.1.6) - barley
C;Species: Hordeum vulgare (barley)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 05-May-2000
C;Accession: S52015; S41949
R;Gregersen, P.L.; Christensen, A.B.; Sommer-Knudsen, J.; Collinge, D.B.
Plant Mol. Biol. 26, 1797-1806, 1994
A;Title: A putative O-methyltransferase from barley is induced by fungal pathogens and U
A;Reference number: S52015; MUID:95161704; PMID:7858218
A;Accession: S52015
A;Molecule type: mRNA
A;Residues: 1-390 <GRE>
A;Cross-references: EMBL:X77467
R;Gregersen, P.L.; Christensen, A.B.; Sommer-Knudsen, J.; Collinge, D.B.
submitted to the EMBL Data Library, February 1994
A;Description: A novel putative O-methyltransferase from barley is induced by fungal pat
A;Reference number: S41949
A;Accession: S41949
A;Molecule type: mRNA
A;Residues: 1-315,'H',317-390 <GRW>
A;Cross-references: EMBL:X77467; NID:g453243; PIDN:CAA54616.1; PID:g453244
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase; S-adenosylmethionine

  Query Match             34.6%;  Score 652.5;  DB 2;  Length 390;
  Best Local Similarity   39.1%;  Pred. No. 1.2e-44;
  Matches  144;  Conservative   67;  Mismatches  134;  Indels   23;  Gaps    7;

Qy          6 NNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQ 65
              :||    :| | :| |: |  |:: |  |: :|:| |||::  | :|   :  ||:| :
Db         31 SNG----QILQAEAELFCHSFGYLKSMALQSVVKLRIPDVL--HRYGGAASLPELLSTVP 84

Qy         66 VPPTKTRQVQSLMRYLAHNGFFEIVRIHDNI------EAYALTAASELLVKSSEL----S 115
              : | |   :  ||: || | |    :  :        | | | | |||  : :    |
Db         85 IHPNKLPYLPRLMKMLAAAGIFTAEDVPATVGDGEPTTLYHLNAVSRLLVDDASVNGGAS 144

Qy        116 LAPMVEYFLEPNCQGAWNQLKRWVHEEDL----TVFEVSLGTPFWDFINKDPAYNKSFNE 171
              ::| |     |  ||  :|  |:  |:     | | :: |   :   :|  :|  ||:
Db        145 MSPCVLLGTVPLFLGASLKLHEWLQSEEQATTETPFMLAHGGTLYGIGGRDSEFNTVFNK 204

Qy        172 AMACDSQML-NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVE 230
              ||   |: :  || |:|  || |::|:||| || | ||: | |||| :|| ||: | |::
Db        205 AMGASSEFVAALAVRECRDVFAGIKSLVDVAGGNGTTARTIAEAFPYVKCSVLDLPQVIQ 264

Qy        231 NLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVV 290
              :|    : || ||| : :| |:|||| |||||:| ||:|||   |:||||    | |||:
Db        265 GISSHGTVEFVAGDMMEFVPPAEAVLLKYVLHNWSDQDCVKILTRCREAISPGEKAGKVI 324

Qy        291 VIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGY 350
              :||||:     :|: | :: ||: |  : ||| |:|::| |:|:|||| |||     |
Db        325 IIDTVV--GSPSQQILESQVTMDLSMMMLFNGKVREEQNWHKIFLEAGFSHYKIHNVLGM 382

Qy        351 LSLIEIYP 358
              ||||: |
Db        383 RSLIEVQP 390


RESULT 6
```

```
JQ2268
O-methyltransferase (EC 2.1.1.-) - maize
C;Species: Zea mays (maize)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 29-Sep-1999
C;Accession: JQ2268
R;Held, B.M.; Wang, H.; John, I.; Wurtele, E.S.; Colbert, J.T.
Plant Physiol. 102, 1001-1008, 1993
A;Title: An mRNA putatively coding for an O-methyltransferase accumulates preferential
A;Reference number: JQ2268; MUID:94105316; PMID:8278520
A;Accession: JQ2268
A;Molecule type: mRNA
A;Residues: 1-364 <HEL>
A;Cross-references: GB:L14063; NID:g404069; PIDN:AAA18532.1; PID:g404070
A;Experimental source: root, cv. NKH31
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

  Query Match             32.1%;  Score 606;  DB 2;  Length 364;
  Best Local Similarity   35.9%;  Pred. No. 5.8e-41;
  Matches  132;  Conservative   74;  Mismatches  148;  Indels   14;  Gaps    7;

Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              |  | ||    | :  |  |:     |: |  ||  : | | | |  | ||   : |::
Db          1 MELSPNNSTDQS-LLDAQLELWHTTFAPMKSMALKSAIHLRIADAI--HLHGGAASLSQI 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRI-----HDNIEAYALTAASELLV--KSSE 113
              :| : : |::   :: ||| |  |    :     |:   | ||  | ||:  :||:
Db         58 LSKVHLHPSRVSSLRRLMRVLTTTNVFGTQPLGGGSDDDSEPVYTLTPVSRLLIGSQSSQ 117

Qy        114 LSLAPMVEYFLEPNCQGAWNQLKRWVHEE--DLTVFEVSLGTPFWDFINKDPAYNKSFNE 171
              |:  |:    |:|     :::|  |   |  | :|: : |   |:    |  ::   |:
Db        118 LAQTPLAAMVLDPTIVSPFSELGAWFQHELPDPCIFKHTHGRGIWELTKDDATFDALVND 177

Qy        172 AMACDSQML-NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVE 230
              :| |||:: ::| :    ||:|: |:|||||  |   |: | :||| :|| ||:  :||
Db        178 GLASDSQLIVDVAIKQSAEVFQGISSLVDVGGGIGAAAQAISKAFPHVKCSVLDLAHVVA 237

Qy        231 NLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVV 290
                  :: |: |||: || ||||||| |||:|: :||:|||:|||:||       |||:
Db        238 KAPTHTDVQFIAGDMFESIPPADAVLLKSVLHDWDHDDCVKILKNCKKAIPPREAGGKVI 297

Qy        291 VIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGY 350
              :|: |:     : :  |:: : ||::   ||| || |::| |:| |||:  |:| |  |
Db        298 IINMVVGAGPSDMKHKEMQAIFDVYI-MFINGMERDEQEWSKIFSEAGYSDYRIIPVLGV 356

Qy        351 LSLIEIYP 358
              |:||:||
Db        357 RSIIEVYP 364


RESULT 7
T09600
catechol O-methyltransferase homolog - Monterey pine
C;Species: Pinus radiata (Monterey pine)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
C;Accession: T09600
R;Wagner, A.; Walden, A.; Narayan, R.; Walter, C.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z16762
A;Accession: T09600
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-382 <WAG>
A;Cross-references: EMBL:U70873; NID:g1568663; PID:g1568664
A;Experimental source: tissue-type male cone
C;Genetics:
A;Gene: omt
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

  Query Match             28.1%;  Score 529.5;  DB 2;  Length 382;
  Best Local Similarity   33.5%;  Pred. No. 8.6e-35;
  Matches  131;  Conservative   69;  Mismatches  148;  Indels   43;  Gaps   11;
```

```
Qy          1 MASSLN-----NGRKAS---------EIFQGQALLYKHLLGFIDSKCLKWMVELDIPDII 46
              | |::|     || : |         |  ||||  :|    | :|  :| :| | |||:|
Db          1 MDSNMNGLAKSNGCEISRDGFFESEEEELQGQAEAWKCTFAFAESLAVKCVVLLGIPDMI 60

Qy         47 HSHSHGQPITFSELVSILQVPPTKTRQVQSL---MRYLAHNGFF---EIVRIHDNIEA-Y 99
                ::  |:|: |   ||::     |   ||:|  | |   :  |     | |
Db         61 AREGSRATLSLGEIVAKL---PTESPDAACLFRIMRFLVAKGIFRASKSAREGGAFETRY 117

Qy        100 ALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFI 159
               || ||: |||  |||:|||:    :      |:     | |  :  |: : |   | :
Db        118 GLTPASKWLVKGRELSMAPMLLMQNDETTLAPWHHFNECVLEGGV-AFQKANGAEIWSYA 176

Qy        160 NKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLK 219
              :  | :|  || |||:::::         |  | |:|||||||     |  |:| :
Db        177 SDHPDFNNLFNNAMACNARIVMKGILSKYQGFHSLNSLVDVGGGTGTAVAEIVRAYPFIT 236

Qy        220 CMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEA 279
              :  : |:||   |  : :  ||||||: :|  ||: :| ::|:||| ||:|||:||::|
Db        237 GINYDLPHVVATASSLSGVQHVGGDMFETVPTGDAIFMKWIMHDWNDEDCIKILKNCRKA 296

Qy        280 ISGESKTGKVVVIDTVIN----ENKDERQV--------TELKLLMDVHMACIINGKERKE 327
              |     ||||:::| |::    :| |:::          |   |:|  | :   ||||
Db        297 I---PDTGKVIIVDVVLDADQGDNTDKKRKKAVDPIVGTVFDLVMVAHSS---GGKERSE 350

Qy        328 EDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              ::||:: :| ||  | |        |:|| :|
Db        351 KEWKRILLEGGFSRYNIIEIPALQSVIEAFP 381
```

RESULT 8
T04963
catechol O-methyltransferase homolog T12J5.30 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 26-Aug-1999
C;Accession: T04963
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15392
A;Accession: T04963
A;Molecule type: DNA
A;Residues: 1-382 <BEV>
A;Cross-references: EMBL:AL035522
A;Experimental source: cultivar Columbia; BAC clone T12J5
C;Genetics:
A;Map position: 4
A;Introns: 275/3
A;Note: T12J5.30
C;Superfamily: O-methyltransferase

Query Match 27.9%; Score 527; DB 2; Length 382;
Best Local Similarity 34.1%; Pred. No. 1.4e-34;
Matches 126; Conservative 70; Mismatches 151; Indels 22; Gaps 10;

```
Qy          6 NNGRKA-SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSIL 64
              :| ||    |  :    ::|:: || |     |  ::| ||: | :|   ||:| :|| | :
Db         17 DNKRKVLDEEAKASLDIWKYVFGFADIAAAKCAIDLKIPEAIENHPSSQPVTLAELSSAV 76

Qy         65 QVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNI-EAYALTAASE--LLVKSSELSLAPMVE 121
                |:  |::   ||:| | | |: :   | :   |  | |   :: :    ||||  |
Db         77 SASPSHLRRI---MRFLVHQGIFKEIPTKDGLATGYVNTPLSRRLMITRRDGKSLAPFVL 133

Qy        122 YFLEPNCQGAWNQLKRW----VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDS 177
              :   |     | :|       |:      |:   |  | |   :|  :   |||||||:
Db        134 FETTPEMLAPWLRLSSVVSSPVNGSTPPPFDAVHGKDVWSFAQDNPFLSDMINEAMACDA 193

Qy        178 -QMLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSN 236
               :::      |: :|:|: ::|||||||| |  :: : || :|    : |:|:|     :
Db        194 RRVVPRVAGACHGLFDGVTTMVDVGGGTGETMGMLVKEFPWIKGFNFDLPHVIEVAEVLD 253

Qy        237 NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI 296
              :  | ||||  ||  ||: :| |||:| | ||:|||:|||||:      |||:::::||
Db        254 GVENVEGDMFDSIPACDAIFIKWVLHDWGDKDCIKILKNCKEAV--PPNIGKVLIVESVI 311
```

```
Qy        297 NENK-----DER--QVTELKLLMD-VHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
              |||      |||  ::  ::|::| | ||    ||||  ::|  :  ||||  |::
Db        312 GENKKTMIVDERDEKLEHVRLMLDMVMMAHTSTGKERTLKEWDFVLKEAGFARYEVRDID 371

Qy        349 GYLSLIEIY 357
                |||  |
Db        372 DVQSLIIAY 380
```

RESULT 9
T04962
catechol O-methyltransferase homolog T12J5.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 24-Nov-1999
C;Accession: T04962
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15392
A;Accession: T04962
A;Molecule type: DNA
A;Residues: 1-325 <BEV>
A;Cross-references: EMBL:AL035522
A;Experimental source: cultivar Columbia; BAC clone T12J5
C;Genetics:
A;Map position: 4
A;Introns: 109/1; 218/3
A;Note: T12J5.20
C;Superfamily: O-methyltransferase

Query Match 25.7%; Score 485; DB 2; Length 325;
Best Local Similarity 32.3%; Pred. No. 2.6e-31;
Matches 113; Conservative 64; Mismatches 115; Indels 58; Gaps 8;

```
Qy         21 LYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSLMRY 80
              ::::: || |    |  ::| ||: | :|   ||:| ||| | :   |:  |::   ||:
Db         19 IWRYVFGFADIAAAKCAIDLKIPEAIENHPSSQPVTLSELSSAVSASPSHLRRI---MRF 75

Qy         81 LAHNGFFEIVRIHDNI-EAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQLKRWV 139
              | | | |: |  | :  | | |  :: :                          :
Db         76 LVHQGLFKEVPTKDGLATGYTNTPLSRRMMITK-------------------------L 109

Qy        140 HEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDS-QMLNLAFRDCNWVFEGLESIV 198
              | :||           | |   :  :::  |||||||: :::       |  :|:|: ::|
Db        110 HGKDL-----------WAFAQDNLCHSQLINEAMACDARRVVPRVAGACQGLFDGVATVV 158

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
              ||||||| '  |: : || :|    : |:|:|      : :  | ||||  || :|||::|
Db        159 DVGGGTGETMGILVKEFPWIKGFNFDLPHVIEVAQVLDGVENVEGDMFDSIPASDAVIIK 218

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD-----------ERQVTE 307
              |||:| | ||:|||:|||||:      |||:::: || | |:           |    :
Db        219 WVLHDWGDKDCIKILKNCKEAVL--PNIGKVLIVECVIGEKKNTMIAEERDDKLEHVRLQ 276

Qy        308 LKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIY 357
              | ::| || :    ||||  ::|  :  ||||  |::  |    |||  |
Db        277 LDMVMMVHTS---TGKERTLKEWDFVLTEAGFARYEVRDFDDVQSLIIAY 323
```

RESULT 10
T01354
herbicide safener binding protein 1 - maize
C;Species: Zea mays (maize)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Oct-1999
C;Accession: T01354
R;Scott-Craig, J.S.; Casida, J.E.; Poduje, L.; Walton, J.D.
Plant Physiol. 116, 1083-1089, 1998
A;Title: Herbicide safener-binding protein of maize. Purification, cloning, and expres
A;Reference number: Z14299; MUID:98169422; PMID:9501141
A;Accession: T01354
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-363 <SCO>

```
A;Cross-references: EMBL:AF033496; NID:g2921303; PIDN:AAC12715.1; PID:g2921304
A;Experimental source: strain B73
C;Genetics:
A;Gene: SBP1
A;Map position: 2
C;Superfamily: O-methyltransferase

  Query Match             22.6%;  Score 426;  DB 2;  Length 363;
  Best Local Similarity   31.4%;  Pred. No. 1.6e-26;
  Matches  116;  Conservative   70;  Mismatches  144;  Indels   40;  Gaps   10;

Qy         12 SEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKT 71
Db         11 AELLKAQADIWRLSLSYLTPLSLRCAVELGIPTAI--YRHGGAASAAELVTALSLPSTKL 68

Qy         72 RQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLV---------KSSELSLAPMVEY 122
Db         69 PFLRRLLRLLAASGVFTVDKQSSEEERYRISPVSYLLVDGIPHEDHMNHTALVLTCTSTR 128

Qy        123 FLEPNCQGAWNQLKRWVHE-------EDL---TVFEVSLGTPFWDFINKDPAYNKSFNEA 172
Db        129 YIE-----AGIGLAEWFKRDVVTSPFEELHGATLFHESMG-------SLDADFHDMASEA 176

Qy        173 M-ACDSQMLNLAFRDCNWVFEGLESIVDVGGGTGIT--AKIICEAFPKLKCMVLERPNVV 229
Db        177 LDAHDNFGIEIAMREFRDLFEGIQSMTYCCGNFGDDKGARAIVKAFPHIKCTVLAPPKII 236

Qy        230 ENLSGSN-NLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGK 288
Db        237 ATKPADGAMINYVEGDMFSFIPPAQTVVLKLVLHHLTDEECVKLLAQCRKAIPSRKDGGK 296

Qy        289 VVVIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFT 348
Db        297 VIIGDIMIDHS---GPMLETHLLMDIGMMTMTKGRQRDEKEWSELFTKAGFSEYKILKEF 353

Qy        349 GYLSLIEIYP 358
Db        354 GARVAFEVYP 363


RESULT 11
S36403
catechol O-methyltransferase (EC 2.1.1.6) - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 05-May-2000
C;Accession: S36403
R;Legrand, M.
submitted to the EMBL Data Library, August 1993
A;Reference number: S36403
A;Accession: S36403
A;Molecule type: mRNA
A;Residues: 1-364 <LEG>
A;Cross-references: EMBL:X74452; NID:g396588; PID:g396589
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase; S-adenosylmethionine

  Query Match             20.7%;  Score 390.5;  DB 2;  Length 364;
  Best Local Similarity   30.4%;  Pred. No. 1.2e-23;
  Matches  103;  Conservative   64;  Mismatches  135;  Indels   37;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL-----QVPPTKTRQVQSLMRYLAHNGFFE 88
Db         36 LKSAVELDLLELMAKAGPGAAISPSELAAQLSTQNPEAPVMLDRMLRLLASYSVLNCTLR 95

Qy         89 IVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHEED 143
Db         96 TLPDSSVERLYSLAPVCKYLTKNADGVSVAPLL----------LMNQDKVLMESWYHLKD 145

Qy        144 LTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
Db        146 AVLDGGIPFNKAYGMTAFEYHGTDPRFNKVFNRGMSDHSTMSMKKILEDYKGFEGLNSIV 205

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
Db        206 DVGGGTGATVNMIVSKYPSIKGINFDLPHVIGDAPTYPGVEHVGGDMFASVPKADAIFMK 265

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
Db        266 WICHDWSDEHCLKFLKNCYEALPA---NGKVIIAECILPEAPDTSLATKNTVHVDIVMLA 322

Qy        318 CIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEI 356
Db        323 HNPGGKERTEKEFEALAKGAG--------FTGFARLVAL 353


RESULT 12
S36404
catechol O-methyltransferase (EC 2.1.1.6) - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 05-May-2000
C;Accession: S36404
R;Legrand, M.
submitted to the EMBL Data Library, August 1993
A;Reference number: S36403
A;Accession: S36404
A;Molecule type: mRNA
A;Residues: 1-364 <LEG>
A;Cross-references: EMBL:X74453; NID:g396590; PID:g396591
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase; S-adenosylmethionine

  Query Match             20.5%;  Score 386.5;  DB 2;  Length 364;
  Best Local Similarity   30.1%;  Pred. No. 2.4e-23;
  Matches  102;  Conservative   64;  Mismatches  136;  Indels   37;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSIL-----QVPPTKTRQVQSLMRYLAHNGFFE 88
Db         36 LKSALELDLLELMAKAGPGAAISPSELAAQLSTQNPEAPVILDRMLRLLATYSVLNCTLR 95

Qy         89 IVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHEED 143
Db         96 TLSDGSVERLYSLAPVCKFLTKNADGVSVAPLL----------LMNQDKVLMESWYHLKD 145

Qy        144 LTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
Db        146 AVLDGGIPFNKAYGMTAFEYHGTDPRFNKVFNRGMSDHSTMSMKKILEDYKGFEGLNSIV 205

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
Db        206 DVGGGTGATVNMIVSKHPSIKGINFDLPHVIGDAPAYPGVEHVGGDMFASVPKADAIFMK 265

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
Db        266 WICHDWSDEHCLKFLKNCYEALPA---NGKVIIAECILPEAPDTSLATKNTVHVDIVMLA 322

Qy        318 CIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEI 356
Db        323 HNPGGKERTEKEFEALAKGAG--------FTGFARLVAL 353


RESULT 13
T09617
isoliquiritigenin 2'-O-methyltransferase - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T09617
R;Maxwell, C.A.; Harrison, M.J.; Dixon, R.A.
Plant J. 4, 971-981, 1993
A;Title: Molecular characterization and expression of alfalfa isoliquiritigenin 2'-O-m
ion genes.
A;Reference number: Z16778; MUID:94108491; PMID:8281189
A;Accession: T09617
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-372 <MAX>
A;Cross-references: EMBL:L10211; NID:g289126; PIDN:AAB48059.1; PID:g1843462
```

```
A;Experimental source: subspecies sativa; cultivar Apollo
C;Function:
A;Description: methylates the 2'-hydroxyl of isoliquiritigenin (2',4,4'-trihydroxychalco
A;Note: 4,4'-dihydroxy-2'-methoxychalcone is the most potent of the nod-gene-inducing fl
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

  Query Match           20.4%;  Score 384;  DB 2;  Length 372;
  Best Local Similarity  28.5%;  Pred. No. 4e-23;
  Matches  103;  Conservative   77;  Mismatches  137;  Indels   44;  Gaps   12;

Qy         30 DSKCLKWMV--------------ELDIPDII-HSHSHGQPITFSELVSILQVPPTKTR- 72
              || ||  ||                :|:: :||  :   |  :: ||: | |   |  |:
Db         22 DSACLSAMVLTTNLVYPAVLNAAIDLNLFEIIAKATPPGAFMSPSEIASKL---PASTQH 78

Qy         73 -----QVQSLMRYLAHNGFFEIVRIHDNIE------AYALTAASELLV-KSSELSLAPMV 120
                  ::  ::| ||   :  :       ||        | |:   : ||   |   ||
Db         79 SDLPNRLDRMLRLLA--SYSVLTSTTRTIEDGGAERVYGLSMVGKYLVPDESRGYLASFT 136

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAM--ACDSQ 178
               :   |      |   |  | :||: :|:   |   ::|: ||   |: ||::|   | ::
Db        137 TFLCYPALLQVWMNFKEAVVDEDIDLFKNVHGVTKYEFMGKDKKMNQIFNKSMVDVCATE 196

Qy        179 MLNLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNL 238
              |  :   :    |||: ::|||||:|   ::|   :| :| :  :  | |:||    : :
Db        197 MKRML--EIYTGFEGISTLVDVGGGSGRNLELIISKYPLIKGINFDLPQVIENAPPLSGI 254

Qy        239 TFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINE 298
              ||||||  :|: ||::|| | |||:|   |:: | || :|:    |  |||:::: :: |
Db        255 EHVGGDMFASVPQGDAMILKAVCHNWSDEKCIEFLSNCHKAL---SPNGKVIIVEFILPE 311

Qy        299 NKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKIS--PFTGYLSLIEI 356
                :  : ::|   :|  |   : |:|| |: ::||   :||  ::::   |   | ::|
Db        312 EPNTSEESKLVSTLDNLMFITVGGRERTEKQYEKLSKLSGFSKFQVACRAFNS-LGVMEF 370

Qy        357 Y 357
              |
Db        371 Y 371


RESULT 14
T12259
O-diphenol-O-methyltransferase (EC 2.1.1.-) - pepper
C;Species: Capsicum annuum (pepper)
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 11-Jan-2000
C;Accession: T12259
R;Lee, B.; Choi, D.; Lee, K.W.
J. Plant Biol. 41, 9-14, 1998
A;Title: Isolation and characterization of o-diphenol-O-methyltransferase cDNA Clone in
A;Reference number: Z17476
A;Accession: T12259
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-359 <LEE>
A;Cross-references: EMBL:U83789; NID:g1791351; PID:g1791352
A;Experimental source: very young green fruit
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase

  Query Match           20.3%;  Score 383.5;  DB 2;  Length 359;
  Best Local Similarity  31.3%;  Pred. No. 4.1e-23;
  Matches  103;  Conservative   60;  Mismatches  131;  Indels   35;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKT--------RQVQSLMRYLAHNG 85
              ||  :|||: :|:      |  |: ||| : |   |||          | :: |  |    |
Db         32 LKSALELDLLEIMAKAGPGAAISPSELAAQL---PTKNPEAPVMLDRMLRLLATYSVLNC 88

Qy         86 FFEIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVH 140
                    :       |:|    :|| |::: :|:||::             ||    :: | |
Db         89 TLRTLPDGRVERLYSLAPVCKLLTKNADGVSVAPLL----------LMNQDKVLMESWYH 138

Qy        141 EEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLE 195
                 |  :     |  : |   :::   || :|| ||  |:  | |      :    ||||

Db        139 LTDAVLDGGVPFNKAYGMTAFEYHGTDPRFNKVFNRGMSDHSTMTMKKILEDYKGPEGLN 198

Qy        196 SIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAV 255
              |||||||||| |  :|   :| :| :  :  :|:|:       :  || |||  :|||||:
Db        199 SIVDVGGGTGATVNMIVSKYPSIKGINFDLSHVIEDAPAYPGVEHVGRDMFVSVPKADAI 258

Qy        256 LLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVH 315
               :| : |:|:|  |:| |:|| ||:       |||:| : :: |  |     |:  : :|:
Db        259 FMKWICHDWSDEHCLKFLKNCYEALPA---NGKVLVAECILPETPDTSAATKNAVHVDIV 315

Qy        316 M-ACIINGKERKEEDWKKLFMEAGFQSYK 343
              | |    |||| |:::: |   |||  ::
Db        316 MLAHNPGGKERTEKEFEALAKGAGFTGFR 344


RESULT 15
T09673
caffeate O-methyltransferase (EC 2.1.1.68) - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T09673
R;Gowri, G.; Bugos, R.C.; Campbell, W.H.; Maxwell, C.A.; Dixon, R.A.
Plant Physiol. 97, 7-14, 1991
A;Title: Molecular cloning and expression of alfalfa S-adenosyl-L-methionine: caffeic
A;Reference number: Z16815
A;Accession: T09673
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <GOW>
A;Cross-references: EMBL:M63853; NID:g166419; PID:g166420
A;Experimental source: subspecies sativa, cultivar Apollo
C;Function:
A;Pathway: lignin biosynthesis
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase; S-adenosylmethionine

  Query Match           20.1%;  Score 379.5;  DB 2;  Length 365;
  Best Local Similarity  31.4%;  Pred. No. 8.9e-23;
  Matches  106;  Conservative   63;  Mismatches  134;  Indels   35;  Gaps   10;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSLMRYLAHNGFFEI---- 89
              ||  :|||: :||      |  |:  |: |  |:| |        | | |      : |
Db         36 LKSALELDLLEIIAKAGPGAQISPIEIAS--QLPTTNPDAPVMLDRMLRLLACYIILTCS 93

Qy         90 VRIHDNIEA---YALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHE 141
              ||   : :    | |    :: |||: : :|:: :                ||    :: | |
Db         94 VRTQQDGKVQRLYGLATVAKYLVKNEDGVSISAL----------NLMNQDKVLMESWYHL 143

Qy        142 EDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLES 196
              :|  :      |  : |   :::   || :|| ||: |:  | :        :    ||||:|
Db        144 KDAVLDGGIPFNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYTGFEGLKS 203

Qy        197 IVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVL 256
              :||||||||      |   :| :| :  :  |:|:|:       :  ||||||  |||||||
Db        204 LVDVGGGTGAVINTIVSKYPTIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSIPKADAVF 263

Qy        257 LKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM 316
              :| : |:|:|  |:| |:|| ||:       |||:| : ::      |     |:  : :|| |
Db        264 MKWICHDWSDEHCLKFLKNCYEAL---PDNGKVIVAECILPVAPDSSLATKGVVHIDVIM 320

Qy        317 -ACIINGKERKEEDWKKLFMEAGFQSYKI--SPFTGYL 351
               |     |||| ::::: |    |||| :|:  : |  |:
Db        321 LAHNPGGKERTQKEFEDLAKGAGFQGFKVHCNAFNTYI 358


Search completed: November 10, 2003, 03:22:07
Job time : 45 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 10, 2003, 01:53:20 ; Search time 23 Seconds
(without alignments)
731.981 Million cell updates/sec

Title: US-09-868-547-4
Perfect score: 1886
Sequence: 1 MASSLNNGRKASEIFQGQAL.....FQSYKISPFTGYLSLIEIYP 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	636.5	33.7	347	1	6OMT_COPJA	Q9le16 coptis japo
2	606	32.1	364	1	ZRP4_MAIZE	P47917 zea mays (m
3	587.5	31.2	350	1	4OMT_COPJA	Q9le15 coptis japo
4	390.5	20.7	350	1	COMT_COFCA	Q8l187 coffea cane
5	390.5	20.7	363	1	OMT1_ARATH	Q9fk25 arabidopsis
6	383.5	20.3	359	1	COMT_CAPAN	Q9fqy8 capsicum an
7	383.5	20.3	363	1	COMT_CATRO	Q8w013 catharanthu
8	380	20.1	343	1	OMT1_CHRAE	P59049 chrysosplen
9	379.5	20.1	365	1	COMT_MEDSA	P28002 medicago sa
10	379	20.1	343	1	OMT2_CHRAE	Q42653 chrysosplen
11	375.5	19.9	359	1	COMT_CAPCH	O81646 capsicum ch
12	371.5	19.7	361	1	COM1_OCIBA	Q9xgw0 ocimum basi
13	371.5	19.7	365	1	COM1_POPTM	Q00763 populus tre
14	368.5	19.5	365	1	COM1_POPKI	Q43046 populus kit
15	367.5	19.5	365	1	COMT_PRUDU	Q43609 prunus dulc
16	366.5	19.4	364	1	COM3_POPKI	Q43047 populus kit
17	363.5	19.3	361	1	COM2_OCIBA	Q9xgv9 ocimum basi
18	362.5	19.2	364	1	COM2_POPTM	Q41086 populus tre
19	358.5	19.0	366	1	COMT_EUCGU	P46484 eucalyptus
20	356.5	18.9	365	1	COMT_ROSCH	Q8gu25 rosa chinen
21	356	18.9	370	1	COMT_CLABR	O23760 clarkia bre
22	345	18.3	354	1	COMT_ZINEL	Q43239 zinnia eleg
23	339.5	18.0	362	1	COMT_SACOF	O82054 saccharum o
24	334.5	17.7	313	1	COMT_EUCGL	Q9swc2 eucalyptus
25	332.5	17.6	381	1	SMT_COPJA	Q39522 coptis japo
26	327.5	17.4	364	1	COMT_MAIZE	Q06509 zea mays (m
27	303	16.1	365	1	IMT1_MESCR	P45986 mesembryant
28	268	14.2	355	1	CM4T_STRPE	Q06528 streptomyce
29	252	13.4	376	1	DMPM_STRLP	P42712 streptomyce
30	202	10.7	494	1	TCMN_STRGA	P16559 streptomyce
31	194	10.3	345	1	HIOM_BOVIN	P10950 bos taurus
32	194	10.3	345	1	HIOM_HUMAN	P46597 homo sapien
33	192	10.2	346	1	HIOM_CHICK	Q92056 gallus gall
34	188.5	10.0	621	1	ASML_HUMAN	O95671 homo sapien
35	188	10.0	418	1	OMTA_ASPFL	P55790 aspergillus
36	187	9.9	418	1	OMTA_ASPPA	Q12120 aspergillus
37	151	8.0	376	1	CRTF_RHOSH	P54906 rhodobacter
38	139.5	7.4	393	1	CRTF_RHOCA	P17061 rhodobacter
39	136.5	7.2	339	1	TCMO_STRGA	P39896 streptomyce
40	111	5.9	4568	1	DYHC_CAEEL	Q19020 caenorhabdi
41	98	5.2	541	1	UME5_YEAST	P39073 saccharomyc
42	98	5.2	1037	1	N120_YEAST	P35729 saccharomyc
43	97.5	5.2	983	1	EPA3_HUMAN	P29320 homo sapien
44	97	5.1	944	1	SYL2_SULSO	O33768 sulfolobus
45	96.5	5.1	459	1	TRME_STAAM	Q931e1 staphylococ

ALIGNMENTS

RESULT 1
6OMT_COPJA
ID 6OMT_COPJA STANDARD; PRT; 347 AA.
AC Q9LEL6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE (R,S)-norcoclaurine 6-O-methyltransferase (EC 2.1.1.128) (S-adenosyl-
DE L-methionine:norcoclaurine 6-O-methyltransferase) (6-OMT).
OS Coptis japonica (Japanese goldthread).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC Ranunculaceae; Coptis.
OX NCBI_TaxID=3442;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20390108; PubMed=10811648;
RA Morishige T., Tsujita T., Yamada Y., Sato F.;
RT "Molecular characterization of the S-adenosyl-L-methionine:
RT 3'-hydroxy-N-methylcoclaurine 4'O-methyltransferase involved in
RT isoquinoline alkaloid biosynthesis in Coptis japonica.";
RL J. Biol. Chem. 275:23398-23405(2000).
RN [2]
RP SEQUENCE OF 88-103, AND CHARACTERIZATION.
RX MEDLINE=95010097; PubMed=7925429;
RA Sato F., Tsujita T., Katagiri Y., Yoshida S., Yamada Y.;
RT "Purification and characterization of S-adenosyl-L-methionine:
RT norcoclaurine 6-O-methyltransferase from cultured Coptis japonica
RT cells.";
RL Eur. J. Biochem. 225:125-131(1994).
CC -!- FUNCTION: CATALYZES THE TRANSFER OF THE S-METHYL GROUP OF S-
CC ADENOSYL-L-METHIONINE (ADOMET) TO THE 6-HYDROXYL GROUP OF
CC NORCOCLAURINE TO FORM COCLAURINE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + (R,S)-norcoclaurine
CC = S-adenosyl-L-homocysteine + (R,S)-coclaurine.
CC -!- PATHWAY: CARRIES OUT A STEP IN THE CONVERSION OF 3'-HYDROXY-N-
CC METHYLCOCLAURINE TO RETICULINE, AN IMPORTANT INTERMEDIATE IN
CC SYNTHESIZING ISOQUINOLINE ALKALOIDS.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC --

CC --
DR EMBL; D29811; BAB08004.1; -.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR001077; O_Metransf2.
DR Pfam; PF00891; Methyltransf_2; 1.
KW Transferase; Methyltransferase.
SQ SEQUENCE 347 AA; 38700 MW; 2BC34B3FBE67167C CRC64;

```
  Query Match           33.7%;  Score 636.5;  DB 1;  Length 347;
  Best Local Similarity   38.9%;  Pred. No. 5.2e-44;
  Matches  133;  Conservative   70;  Mismatches  132;  Indels    7;  Gaps    5;

Qy         18 QALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSL 77
              || |:  : || :|  ||  |:||: :||  |: |  :| ||| | |   |     :  :
Db         11 QAKLWNFIYGFAESLVLKCAVQLDLANII--HNSGTSMTLSELSSRLPSQPVNEDALYRV 68

Qy         78 MRYLAHNGFPEIVRIHDNIEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQLKR 137
              |||| |   |     |   :  | |   :: ||| :  :  :    : :    |: ||
Db         69 MRYLVHMKLFTKASIDGELR-YGLAPPAKYLVKGWDKCMVGSILAITDKDFMAPWHYLKD 127

Qy        138 WVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAF-RDCNWVFEGLES 196
              :  |  | || :|||  | :: : |  |: ||||||| ||:::  |  ::|  :| |: :
Db        128 GLSGESGTAFEKALGTNIWGYMAEHPEKNQLFNEAMANDSRLIMSALVKECGNIFNGITT 187

Qy        197 IVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVL 256
              :|||||||   : |  ||| :|| | : |:|: :  | : :  | ||||| ||||||::
Db        188 LVDVGGGTGTAVRNIANAFPHIKCTVYDLPHVIADSPGYSEVHCVAGDMFKFIPKADAIM 247

Qy        257 LKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM 316
              :| :||:|:| :|::||: |||| :    | |||:::| |:|  : |   |:::| :|: |
Db        248 MKCILHDWDDKECIEILKRCKEAV--PVKGGKVIIVDIVLNV-QSEHPYTKMRLTLDLDM 304

Qy        317 ACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              |||| ||:|||| :||:: :||:  |   |:|| ||
Db        305 MLNTGGKERTEEEWKKLIHDAGYKGHKITQITAVQSVIEAYP 346
```

```
RESULT 2
ZRP4_MAIZE
ID   ZRP4_MAIZE     STANDARD;      PRT;   364 AA.
AC   P47917;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   O-methyltransferase ZRP4 (EC 2.1.1.-) (OMT).
GN   ZRP4.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NKH31; TISSUE=Root;
RX   MEDLINE=94105316; PubMed=8278520;
RA   Held B.M., Wang H., John I., Wurtele E.S., Colbert J.T.;
RT   "An mRNA putatively coding for an O-methyltransferase accumulates
RT   preferentially in maize roots and is located predominantly in the
RT   region of the endodermis.";
RL   Plant Physiol. 102:1001-1008(1993).
CC   -!- FUNCTION: MAY BE INVOLVED IN THE O-METHYLATION OF SUBERIN
CC       PHENYLPROPANOID PRECURSORS.
CC   -!- TISSUE SPECIFICITY: ACCUMULATES PREFERENTIALLY IN THE ROOTS AND IS
CC       LOCATED PREDOMINANTLY IN THE REGION OF THE ENDODERMIS, LOW LEVELS
CC       ARE SEEN IN THE LEAVES, STEMS, AND OTHER SHOOT ORGANS.
CC   -!- SIMILARITY: TO OTHER OMTS REQUIRING S-ADENOSYL-L-METHIONINE AS
CC       SUBSTRATE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14063; AAA18532.1; -.
DR   PIR; JQ2268; JQ2268.
DR   MaizeDB; 63528; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   364 AA;  39583 MW;  FB8AD93AD5A6611D CRC64;
```

```
  Query Match           32.1%;  Score 606;  DB 1;  Length 364;
  Best Local Similarity   35.9%;  Pred. No. 1.6e-41;
  Matches  132;  Conservative   74;  Mismatches  148;  Indels   14;  Gaps    7;

Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              |  | ||   | :   |  |:    |: |  ||  : | | | |  | ||   : |::
Db          1 MELSPNNSTDQS-LLDAQLELWHTTFAFMKSMALKSAIHLRIADAI--HLHGGAASLSQI 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRI-----HDNIEAYALTAASELLV--KSSE 113
              :| : : |::   :: ||| |     |    :      |:   | ||  | ||:  :||:
Db         58 LSKVHLHPSRVSSLRRLMRVLTTTNVFGTQPLGGGSDDDSEPVYTLTPVSRLLIGSQSSQ 117

Qy        114 LSLAPMVEYFLEPNCQGAWNQLKRWVHEE--DLTVFEVSLGTPFWDFINKDPAYNKSFNE 171
              |:  |:    |:|     :::|  |   |  |  :|: : |   |:    |  ::   |:
Db        118 LAQTPLAAMVLDPTIVSPFSELGAWFQHELPDPCIFKHTHGRGIWELTKDDATFDALVND 177

Qy        172 AMACDSQML-NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVE 230
              :| |||:: ::|  :    ||:|: |:|||||  |  |: | :||| :|| ||:  :||
Db        178 GLASDSQLIVDVAIKQSAEVFQGISSLVDVGGGIGAAAQAISKAFPHVKCSVLDLAHVVA 237

Qy        231 NLSGSNNLTFVGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVV 290
                :: |: ||||: || ||||||| |||:|: :||:|||:|||:||       |||:
Db        238 KAPTHTDVQFIAGDMFESIPPADAVLLKSVLHDWDHDDCVKILKNCKKAIPPREAGGKVI 297

Qy        291 VIDTVINENKDERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGY 350
              :|: |:     : :  |:: : ||::   ||| || |::| |:| |||:  |:| |  |
Db        298 IINMVVGAGPSDMKHKEMQAIPDVYI-MFINGMERDEQEWSKIFSEAGYSDYRIIPVLGV 356

Qy        351 LSLIEIYP 358
              |:||:||
Db        357 RSIIEVYP 364
```

```
RESULT 3
4OMT_COPJA
ID   4OMT_COPJA     STANDARD;      PRT;   350 AA.
AC   Q9LEL5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   3'-hydroxy-N-methyl-(S)-coclaurine 4'-O-methyltransferase
DE   (EC 2.1.1.116) (S-adenosyl-L-methionine:3'-hydroxy-N-methylcoclaurine
DE   4'-O-methyltransferase) (4'-OMT).
OS   Coptis japonica (Japanese goldthread).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC   Ranunculaceae; Coptis.
OX   NCBI_TaxID=3442;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=20390108; PubMed=10811648;
RA   Morishige T., Tsujita T., Yamada Y., Sato F.;
RT   "Molecular characterization of the S-adenosyl-L-methionine:
RT   3'-hydroxy-N-methylcoclaurine 4'O-methyltransferase involved in
RT   isoquinoline alkaloid biosynthesis in Coptis japonica.";
RL   J. Biol. Chem. 275:23398-23405(2000).
CC   -!- FUNCTION: CATALYZES THE TRANSFER OF THE METHYL GROUP TO THE 4'-
CC       HYDROXYL GROUP OF 3'-HYDROXY-N-METHYLCOCLAURINE TO FORM
CC       RETICULINE.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3'-hydroxy-N-methyl-
CC       (S)-coclaurine = S-adenosyl-L-homocysteine + (S)-reticuline.
CC   -!- PATHWAY: CARRIES OUT A STEP IN THE CONVERSION OF 3'-HYDROXY-N-
CC       METHYLCOCLAURINE TO RETICULINE, AN IMPORTANT INTERMEDIATE IN
CC       SYNTHESIZING ISOQUINOLINE ALKALOIDS.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D29812; BAB08005.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   350 AA;  38775 MW;  547835EBCDEF9182 CRC64;

  Query Match             31.2%;  Score 587.5;  DB 1;  Length 350;
  Best Local Similarity   34.3%;  Pred. No. 4.7e-40;
  Matches  119;  Conservative   83;  Mismatches  132;  Indels   13;  Gaps    6;

Qy         16 QGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQ 75
              :  ||  ::|  :  ||  ||   |:  |||  |  |||  :::  ||:  ::|  |  |  |        :
Db         13 KAQAHVWKIIYGFADSLVLRCAVELGIVDIIDNNN--QPMALADLASKLPVSDVNCDNLY 70

Qy         76 SLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMVEYFLEPNCQGAWNQL 135
                ::|||        :  :   |  :  |||    :  ||  ::::  |:  ||:    :  :     |: :
Db         71 RILRYLVKMEILRVEKSDDGQKKYALEPIATLLSRNAKRSMVPMILGMTQKDFMTPWHSM 130

Qy        136 KRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAF----RDCNWVF 191
              |  :  :    |  ||  ::|    |:::    |   ::  |||  ||  ::::|   :      ||    :|
Db        131 KDGLSDNG-TAFEKAMGMTIWEYLEGHPDQSQLFNEGMAGETRLLTSSLISGSRD---MF 186

Qy        192 EGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK 251
              :|::|:||||||  |  |  |  |  :|||  :||  :  :  |:|:  |      |:   :||||||  :|
Db        187 QGIDSLVDVGGGNGTTVKAISDAFPHIKCTLFDLPHVIANSYDLPNIERIGGDMFKSVPS 246

Qy        252 ADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLL 311
              |  |::|||:||:|||  |  :|||:  |:  |:   :    |||:::|   ::|   |   :::   :|:
Db        247 AQAIILKLILHDWNDEDSIKILKQCRNAVPKDG--GKVIIVDVALDEESD-HELSSTRLI 303

Qy        312 MDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :|:  |      ||||  :|  |:|:   |||   ||        |:||::|
Db        304 LDIDMLVNTGGKERTKEVWEKIVKSAGFSGCKIRHIAAIQSVIEVFP 350


RESULT 4
COMT_COFCA
ID   COMT_COFCA     STANDARD;      PRT;   350 AA.
AC   Q8LL87; Q8LL88;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase) (COMT) (CAOMT).
OS   Coffea canephora (Robusta coffee).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Gentianales; Rubiaceae; Ixoroideae; Coffeeae;
OC   Coffea.
OX   NCBI_TaxID=49390;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fruit, and Leaf;
RA   Campa C., Legal L., Khounlotham M., Noirot M., de Kochko A.;
RT   "Complete cDNA sequence of a caffeic acid O-methyltransferase from
RT   Coffea canephora fruit.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF454631; AAN03726.1; -.
DR   EMBL; AF454632; AAN03727.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
FT   CONFLICT    230    230       S -> P (IN REF. 1; AAN03726).
FT   CONFLICT    268    268       K -> R (IN REF. 1; AAN03726).
SQ   SEQUENCE   350 AA;  38258 MW;  D6AC10007BDF8B6B CRC64;

  Query Match             20.7%;  Score 390.5;  DB 1;  Length 350;
  Best Local Similarity   30.8%;  Pred. No. 3.6e-24;
  Matches  101;  Conservative   72;  Mismatches  122;  Indels   33;  Gaps   11;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVP---PTKTRQVQSLMRYLAHNGFFE-- 88
              ||  :|||: ::|     |  :: |||  :  |:|     |       :  ::| ||     :
Db         23 LKSAIELDLLELIAKAGPGAYVSPSELAA--QLPTHNPEAPIMLDRILRLLATYSVLDCK 80

Qy         89 IVRIHD-NIE-AYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHE 141
              :  : |  :|  | |    : | |::: :|:||::           ||    :: | |
Db         81 LNNLADGGVERLYGLAPVCKFLTKNADGVSMAPLL----------LMNQDKVLMESWYHL 130

Qy        142 EDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLES 196
              :|  :      |  : |   :::   || :|| ||: |:  | :       :    ||||::
Db        131 KDAVLDGGIPFNKAYGMTAFEYHGTDPRFNKVFNQGMSNHSTITMKKILEVYRGFEGLKT 190

Qy        197 IVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVL 256
              :|||||||  |  :|   :| :| :  | |:|||:    : :  ||||||  :|| ||:
Db        191 VVDVGGGTGATLNMIISKYPTIKGINFELPHVVEDAPSHSGVEHVGGDMFVSVPKGDAIF 250

Qy        257 LKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM 316
              :| : |:|:|: | |:|:|| :|:       |||:: : |: |  |     |:  : :|| |
Db        251 MKWICHDWSDDHCRKLLKNCYQAL---PDNGKVILAECVLPEAPDTSLATQNVVHVDVVM 307

Qy        317 -ACIINGKERKEEDWKKLFMEAGFQSYK 343
               |    |||| |:::: |   |||: ::
Db        308 LAHNPGGKERTEKEFEALAKGAGFKEFR 335


RESULT 5
OMT1_ARATH
ID   OMT1_ARATH     STANDARD;      PRT;   363 AA.
AC   Q9FK25; O49964; Q42170;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Quercetin 3-O-methyltransferase 1 (EC 2.1.1.76) (AtOMT1) (Flavonol 3-
DE   O-methyltransferase 1) (Caffeic acid/5-hydroxyferulic acid O-
DE   methyltransferase).
GN   OMT1 OR AT5G54160 OR K18G13.3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. C24; TISSUE=Leaf;
RX   MEDLINE=98007869; PubMed=9349713;
RA   Zhang H., Wang J., Goodman H.M.;
RT   "An Arabidopsis gene encoding a putative 14-3-3-interacting protein,
```

```
RT   caffeic acid/5-hydroxyferulic acid O-methyltransferase.";
RL   Biochim. Biophys. Acta 1353:199-202(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98403884; PubMed=9734815;
RA   Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT   Sequence features of the regions of 1,367,185 bp covered by 19
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:203-216(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT   SSP consortium (Salk/Stanford/PGEC).";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 285-363 FROM N.A.
RC   STRAIN=cv. Columbia; TISSUE=Green siliques;
RA   Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
RL   Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   FUNCTION, AND CHARACTERIZATION.
RX   MEDLINE=20166962; PubMed=10700397;
RA   Muzac I., Wang J., Anzellotti D., Zhang H., Ibrahim R.K.;
RT   "Functional expression of an Arabidopsis cDNA clone encoding a
RT   flavonol 3'-O-methyltransferase and characterization of the gene
RT   product.";
RL   Arch. Biochem. Biophys. 375:385-388(2000).
CC   -!- FUNCTION: Methylates OH residues of flavonoid compounds. Substrate
CC       preference is quercetin > myricetin >> luteolin. Dihydroquercetin
CC       is not a substrate. Has an optimal pH of 7.5.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,5,7,3',4'-
CC       pentahydroxyflavone = S-adenosyl-L-homocysteine + 3-methoxy-
CC       5,7,3',4'-tetrahydroxyflavone.
CC   -!- ENZYME REGULATION: Does not require magnesium. Completely
CC       inhibited by 5 mM of either NiSO4 or p-chloromercuribenzoate
CC       (pCMB).
CC   -!- PATHWAY: Phenylpropanoid pathway.
CC   -!- SUBUNIT: Monomer.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U70424; AAB96879.1; -.
DR   EMBL; AB013387; BAB11578.1; -.
DR   EMBL; AY062837; AAL32915.1; -.
DR   EMBL; AY081565; AAM10127.1; -.
DR   EMBL; Z27062; CAA81580.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Methyltransferase; Transferase.
FT   DOMAIN      195    296       SAM-BINDING.
FT   CONFLICT    229    229       D -> N (IN REF. 1).
FT   CONFLICT    295    295       E -> V (IN REF. 4).
FT   CONFLICT    301    301       T -> S (IN REF. 4).
FT   CONFLICT    348    348       V -> C (IN REF. 4).
SQ   SEQUENCE   363 AA;  39618 MW;  B4380028D89C43DC CRC64;

 Query Match             20.7%;  Score 390.5;  DB 1;  Length 363;
 Best Local Similarity   30.0%;  Pred. No. 3.8e-24;

  Matches  106;  Conservative   69;  Mismatches  119;  Indels   59;  Gaps   12;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIH 93
              ||  :|||: :|:    :| |:: :|: | |   |||  :   ::          |:|:
Db         36 LKSALELDLLEIM--AKNGSPMSPTEIASKL---PTKNPEAPVMLD--------RILRL- 81

Qy         94 DNIEAYALTAASELLVKSSELS---------LAPMVEYFLEPNCQGA-------WNQ--- 134
                 ||: | |   : :||         | |: :| |  |  |          ||
Db         82 -------LTSYSVLTCSNRKLSGDGVERIYGLGPVCKY-LTKNEDGVSIAALCLMNQDKV 133

Qy        135 -LKRWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCN 188
               :: | | :|  :     |  : |   :::   || :|| ||  |:  | :     :
Db        134 LMESWYHLKDAILDGGIPFNKAYGMSAFEYHGTDPRFNKVFNNGMSNHSTITMKKILETY 193

Qy        189 WVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKC 248
               |||| |:|||||| | | |:|   :| || :  : |:|:|:     :  ||||||
Db        194 KGFEGLTSLVDVGGGIGATLKMIVSKYPNLKGINFDLPHVIEDAPSHPGIEHVGGDMFVS 253

Qy        249 IPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTEL 308
              :|| ||: :| : |:|:|   |:| |:|| |::    : |||:: : :: |  |    : |
Db        254 VPKGDAIFMKWICHDWSDEHCVKFLKNCYESL---PEDGKVILAECILPETPD----SSL 306

Qy        309 KLLMDVHMACII-----NGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEI 356
                 ||: ||:       |||| |:::: |   :||:  |:      ::|||:
Db        307 STKQVVHVDCIMLAHNPGGKERTEKEFEALAKASGFKGIKVVCDAFGVNLIEL 359


RESULT 6
COMT_CAPAN
ID   COMT_CAPAN     STANDARD;      PRT;   359 AA.
AC   Q9FQY8; P93088;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase) (COMT) (CAOMT).
GN   COMT.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pericarp;
RA   Lee B.-H., Choi D., Lee K.-W.;
RT   "Isolation and characterization of o-diphenol-O-methyltransferase cDNA
RT   clone in hot pepper (Capsicum annuum L.).";
RL   J. Plant Biol. 41:9-14(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Chungyang; TISSUE=Root;
RA   Kim K.-W., Lee S.-W.;
RT   "Isolation and characterization of caffeic acid O-methyltrasferase
RT   cDNA from Capsicum annuum.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- TISSUE SPECIFICITY: Fruit. Not expressed in leaf.
CC   -!- DEVELOPMENTAL STAGE: Expression increases during fruit development
CC       but decreases during ripening.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U83789; AAC17455.1; -.
DR   EMBL; AF212316; AAG43822.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
FT   CONFLICT     71     71       V -> I (IN REF. 2).
FT   CONFLICT    246    246       R -> G (IN REF. 2).
SQ   SEQUENCE   359 AA;  39432 MW;  70599AED75CE1730 CRC64;

  Query Match             20.3%;  Score 383.5;  DB 1;  Length 359;
  Best Local Similarity   31.3%;  Pred. No. 1.4e-23;
  Matches  103;  Conservative   60;  Mismatches  131;  Indels   35;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKT--------RQVQSLMRYLAHNG 85
              ||  :|||: :|:      |  |: ||| : |   |||         | :: |  |   |
Db         32 LKSALELDLLEIMAKAGPGAAISPSELAAQL---PTKNPEAPVMLDRMLRLLATYSVLNC 88

Qy         86 FPEIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVH 140
                 :      |:|    :|| |::: :|:||::            ||    :: | |
Db         89 TLRTLPDGRVERLYSLAPVCKLLTKNADGVSVAPLL----------LMNQDKVLMESWYH 138

Qy        141 EEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLE 195
                |  :     |  : |   :::   || :|| ||  |:  | |       :    ||||
Db        139 LTDAVLDGGVPFNKAYGMTAFEYHGTDPRFNKVFNRGMSDHSTMTMKKILEDYKGFEGLN 198

Qy        196 SIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAV 255
              |||||||||| |  :|   :| :| :  :  :|:|:      :  || |||  :|||||:
Db        199 SIVDVGGGTGATVNMIVSKYPSIKGINFDLSHVIEDAPAYPGVEHVGRDMFVSVPKADAI 258

Qy        256 LLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVH 315
                :| : |:|:|  |:| |:|| ||:      |||:| : :: |  |    |:  : :|:
Db        259 FMKWICHDWSDEHCLKFLKNCYEALPA---NGKVLVAECILPETPDTSAATKNAVHVDIV 315

Qy        316 M-ACIINGKERKEEDWKKLFMEAGFQSYK 343
              | |    |||| |:::: |   |||  ::
Db        316 MLAHNPGGKERTEKEFEALAKGAGFTGFR 344


RESULT 7
COMT_CATRO
ID   COMT_CATRO     STANDARD;      PRT;   363 AA.
AC   Q8W013;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase) (COMT) (CAOMT).
GN   COMT1.
OS   Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae;
OC   Catharanthus.
OX   NCBI_TaxID=4058;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schroeder G., Wehinger E., Schroeder J.;
RT   "Molecular and biochemical characterization of caffeic acid O-
RT   methyltransferase from Catharanthus roseus.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY028439; AAK20170.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
SQ   SEQUENCE   363 AA;  39785 MW;  700D420F98A52E35 CRC64;

  Query Match             20.3%;  Score 383.5;  DB 1;  Length 363;
  Best Local Similarity   29.3%;  Pred. No. 1.4e-23;
  Matches   98;  Conservative   71;  Mismatches  117;  Indels   49;  Gaps   10;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVP---PTKTRQVQSLMRYLAH------- 83
              ||  :|||: ::|     |  :: ||| :  |:|   |       :  ::| ||
Db         36 LKSAIELDLLELIKKSGPGAYVSPSELAA--QLPTQNPDAPVMLDRILRLLASYSVLNCT 93

Qy         84 -----NGFFEIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ--- 134
                  :|  |          |:|    : | |: : :|:| ::            ||
Db         94 LKDLPDGGIE--------RLYSLAPVCKFLTKNEDGVSMAALL----------LMNQDKV 135

Qy        135 -LKRWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCN 188
               :: | | :|  :      |  : |   :::  ||| :|| ||: |:  | ::     :
Db        136 LMESWYHLKDAVLEGGIPFNKAYGMTAFEYHGKDPRFNKVFNQGMSNHSTIIMKKILEIY 195

Qy        189 WVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKC 248
               |:||:::|||||||| |  :|   :| :| :  : |:|:|:       :  ||||||
Db        196 QGFQGLKTVVDVGGGTGATLNMIVSKYPSIKGINFDLPHVIEDAPSYPGVDHVGGDMFVS 255

Qy        249 IPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTEL 308
              :|| ||: :| : |:|:|  |:| |:|| ||:    : |||:: : :: |  |    |:
Db        256 VPKGDAIFMKWICHDWSDAHCLKFLKNCHEAL---PENGKVILAECLLPEAPDSTLSTQN 312

Qy        309 KLLMDVHM-ACIINGKERKEEDWKKLFMEAGFQSY 342
              : :|| | |    |||| |:::: |   |||: :
Db        313 TVHVDVIMLAHNPGGKERTEKEFEALAKGAGFRGF 347


RESULT 8
OMT1_CHRAE
ID   OMT1_CHRAE     STANDARD;      PRT;   343 AA.
AC   P59049;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Quercetin 3-O-methyltransferase 1 (EC 2.1.1.76) (Flavonol 3-O-
DE   methyltransferase 1).
GN   OMT1.
OS   Chrysosplenium americanum (Golden saxifrage).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Saxifragales; Saxifragaceae; Chrysosplenium.
OX   NCBI_TaxID=36749;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Leaf;
RX   MEDLINE=98181116; PubMed=9514654;
RA   Gauthier A., Gulick P.J., Ibrahim R.K.;
RT   "Characterization of two cDNA clones which encode O-methyltransferases
RT   for the methylation of both flavonoid and phenylpropanoid
RT   compounds.";
```

```
RL   Arch. Biochem. Biophys. 351:243-249(1998).
CC   -!- FUNCTION: Methylates OH residues of flavonoid and phenylpropanoid
CC       compounds and shows a higher affinity for flavonoid than
CC       phenylpropanoid substrates. Substrate preference is quercetin
CC       (flavonoid) > luteolin (flavonoid) > 5-hydroxyferulic acid
CC       (phenylpropanoid) > caffeic acid (phenylpropanoid) >> apigenin
CC       (flavonoid) = kempferol (flavonoid). 3,4-dimethylquercetin is not
CC       a substrate.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,5,7,3',4'-
CC       pentahydroxyflavone = S-adenosyl-L-homocysteine + 3-methoxy-
CC       5,7,3',4'-tetrahydroxyflavone.
CC   -!- PATHWAY: Phenylpropanoid pathway.
CC   -!- MISCELLANEOUS: The Vmax value for flavonoid compounds is two to
CC       threefold higher for OMT1 than for OMT2.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   -!- CAUTION: It is not sure if OMT1 and OMT2 are really encoded by two
CC       different genes or if they represent cloning artifacts.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Methyltransferase; Transferase.
FT   DOMAIN      173    274       SAM-BINDING.
SQ   SEQUENCE   343 AA;  37821 MW;  F43F2D306E18CB1C CRC64;

  Query Match             20.1%;  Score 380;  DB 1;  Length 343;
  Best Local Similarity   30.4%;  Pred. No. 2.4e-23;
  Matches  102;  Conservative   68;  Mismatches  131;  Indels   34;  Gaps   12;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFS--ELVSILQVP-PTKTRQVQSLMRYLA-HNGFFEI 89
              ||  :|||: :||     ||    |  |: | |    |     |  ::| |: ::
Db         17 LKSAIELDLLEII----RGQDTCMSPTEIASHLPTTNPDAPAMVDRILRLLSCYSVVTCS 72

Qy         90 VRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHEEDL 144
              ||  |:   | |    : | |: : :|:| :      |     ||    :: | | :|
Db         73 VRSVDDQRVYGLAPVCKYLTKNQDGVSIAAL--------C--LMNQDKVLMESWYHLKDA 122

Qy        145 TV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVD 199
             :     |  : |   :::   || :|| ||  |:  | :           |:|| |:||
Db        123 VLDGGIPFNKAYGMSSFEYHGTDPRFNKVFNRGMSDHSTITMKKVFQAYQGFQGLTSLVD 182

Qy        200 VGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKL 259
             |||||| |  :|  :| ::|:  : |:|:|:      :  |||||  :|| ||: :|
Db        183 VGGGTGATLTMILSKYPTIRCINFDLPHVIEDAPEYPGIEHVGGDMFVSVPKGDAIFMKW 242

Qy        260 VLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDV-HMAC 318
             : |:|:|  |:|:|:|| :|:      |||:: : :: | |    |:  : :||  :|
Db        243 ICHDWSDEHCLKLLKNCYDAL---PNNGKVILAECILPEVPDSSLATKGVVHIDVITVAH 299

Qy        319 IINGKERKEEDWKKLFMEAGFQSYKI--SPFTGYL 351
                |||| |:::: |   |||| :::  :  |  |:
Db        300 NPGGKERTEKEFEALAKAAGFQGFQVFCNAFNTYI 334


RESULT 9
COMT_MEDSA
ID   COMT_MEDSA     STANDARD;      PRT;   365 AA.
AC   P28002;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase) (COMT) (CAOMT).
OS   Medicago sativa (Alfalfa).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX   NCBI_TaxID=3879;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Apollo;
RA   Gowri G., Bugos R.C., Campbell W.H., Maxwell C.A., Dixon R.A.;
RT   "Stress responses in alfalfa (Medicago sativa L). X. Molecular cloning
RT   and expression of S-adenosyl-L-methionine:caffeic acid
RT   3-O-methyltransferase, a key enzyme of lignin biosynthesis.";
RL   Plant Physiol. 97:7-14(1991).
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- TISSUE SPECIFICITY: MORE ABUNDANT IN ROOTS AND STEMS.
CC   -!- INDUCTION: BY INFECTION, PLANT WOUNDING, OR ELICITOR TREATMENT OF
CC       CELL CULTURES.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M63853; AAB46623.1; -.
DR   PIR; T09673; T09673.
DR   PDB; 1KYW; 28-AUG-02.
DR   PDB; 1KYZ; 28-AUG-02.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase; 3D-structure.
SQ   SEQUENCE   365 AA;  39946 MW;  C14B0D75F979C6B6 CRC64;

  Query Match             20.1%;  Score 379.5;  DB 1;  Length 365;
  Best Local Similarity   31.4%;  Pred. No. 2.9e-23;
  Matches  106;  Conservative   63;  Mismatches  134;  Indels   35;  Gaps   10;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSLMRYLAHNGFFEI---- 89
              ||  :|||: :||     |  |:  |: |  |:| |       | | |     : |
Db         36 LKSALELDLLEIIAKAGPGAQISPIEIAS--QLPTTNPDAPVMLDRMLRLLACYIILTCS 93

Qy         90 VRIHDNIEA---YALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHE 141
              ||   : :    | |   :: |||: : :|:: :                ||    :: | |
Db         94 VRTQQDGKVQRLYGLATVAKYLVKNEDGVSISAL----------NLMNQDKVLMESWYHL 143

Qy        142 EDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLES 196
             :|  :     |  : |   :::   || :|| ||: |:  | :       :    ||||:|
Db        144 KDAVLDGGIPFNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYTGFEGLKS 203

Qy        197 IVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVL 256
             :||||||||      |   :| :| :  :  |:|:|:       :  ||||||  |||||||
Db        204 LVDVGGGTGAVINTIVSKYPTIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSIPKADAVF 263

Qy        257 LKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM 316
             :| : |:|:|  |:|  |:|| ||:      |||:| : ::    |    |:  : :|| |
Db        264 MKWICHDWSDEHCLKFLKNCYEAL---PDNGKVIVAECILPVAPDSSLATKGVVHIDVIM 320

Qy        317 -ACIINGKERKEEDWKKLFMEAGFQSYKI--SPFTGYL 351
              |    |||| ::::: |   |||| :|:  : |  |:
Db        321 LAHNPGGKERTQKEFEDLAKGAGFQGFKVHCNAFNTYI 358


RESULT 10
OMT2_CHRAE
ID   OMT2_CHRAE     STANDARD;      PRT;   343 AA.
AC   Q42653;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
```

```
DE   Quercetin 3-O-methyltransferase 2 (EC 2.1.1.76) (Flavonol 3-O-
DE   methyltransferase 2).
GN   OMT2.
OS   Chrysosplenium americanum (Golden saxifrage).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Saxifragales; Saxifragaceae; Chrysosplenium.
OX   NCBI_TaxID=36749;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Leaf;
RX   MEDLINE=98181116; PubMed=9514654;
RA   Gauthier A., Gulick P.J., Ibrahim R.K.;
RT   "Characterization of two cDNA clones which encode O-methyltransferases
RT   for the methylation of both flavonoid and phenylpropanoid
RT   compounds.";
RL   Arch. Biochem. Biophys. 351:243-249(1998).
CC   -!- FUNCTION: Methylates OH residues of flavonoid and phenylpropanoid
CC       compounds and shows a higher affinity for flavonoid than
CC       phenylpropanoid substrates. Substrate preference is quercetin
CC       (flavonoid) = luteolin (flavonoid) = 5-hydroxyferulic acid
CC       (phenylpropanoid) > caffeic acid (phenylpropanoid) >> apigenin
CC       (flavonoid) = kempferol (flavonoid). 3,4-dimethylquercetin is not
CC       a substrate.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,5,7,3',4'-
CC       pentahydroxyflavone = S-adenosyl-L-homocysteine + 3-methoxy-
CC       5,7,3',4'-tetrahydroxyflavone.
CC   -!- PATHWAY: Phenylpropanoid pathway.
CC   -!- MISCELLANEOUS: The Vmax value for flavonoid compounds is two to
CC       threefold lower for OMT2 than for OMT1.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   -!- CAUTION: It is not sure if OMT1 and OMT2 are really encoded by two
CC       different genes or if they represent cloning artifacts.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U16793; AAA86982.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Methyltransferase; Transferase.
FT   DOMAIN      173    274       SAM-BINDING.
SQ   SEQUENCE   343 AA;  37868 MW;  8D363A98330FDE4F CRC64;

  Query Match             20.1%;  Score 379;  DB 1;  Length 343;
  Best Local Similarity   30.4%;  Pred. No. 2.9e-23;
  Matches  102;  Conservative   68;  Mismatches  131;  Indels   34;  Gaps   12;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFS--ELVSILQVP-PTKTRQVQSLMRYLA-HNGFFEI 89
              ||  :|||: :||     ||    |  |: | |     |     | ::| |: ::
Db         17 LKSAIELDLLEII----RGQDTCMSPTEIASHLPTTNPDAPAMVDRILRLLSCYSVVTCS 72

Qy         90 VRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHEEDL 144
              ||  |:   | |    : | |: : :|:| :       |    ||    :: | | :|
Db         73 VRSVDDQRVYGLAPVCKYLTKNQDGVSIAAL--------C--LMNQDKVLMESWYHLKDA 122

Qy        145 TV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIVD 199
              :     |  : |   :::   || :|| ||  |:  | :           |:|| |:||
Db        123 VLDGGIPFNKAYGMSSFEYHGTDPRFNKVFNRGMSDHSTITMKKVFQTYQGFQGLTSLVD 182

Qy        200 VGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLKL 259
              |||||| |  :|   :| ::|:  : |:|:|:      :  ||||||  :|| ||: :|
Db        183 VGGGTGATLTMILSKYPTIRCINFDLPHVIEDAPEYPGIEHVGGDMFVSVPKGDAIFMKW 242

Qy        260 VLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDV-HMAC 318
              : |:|:|  |:|:|:|| :|:      |||:: : :: |  |     |:  : :||  :|
Db        243 ICHDWSDEHCLKLLKNCYDAL---PNNGKVILAECILPEVPDSSLATKGVVHIDVITVAH 299

Qy        319 IINGKERKEEDWKKLFMEAGFQSYKI--SPFTGYL 351
                 |||| |:::: |   |||| :::  : |  |:
Db        300 NPGGKERTEKEFEALAKAAGFQGFQVFCNAFNTYI 334


RESULT 11
COMT_CAPCH
ID   COMT_CAPCH     STANDARD;      PRT;   359 AA.
AC   O81646;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase) (COMT) (CAOMT).
GN   COMT.
OS   Capsicum chinense (Scotch bonnet) (Bonnet pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=80379;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Habanero;
RA   Curry J., Mendoza M., O'Connell M.;
RT   "Nucleotide sequence of a caffeic acid 3-O-methyltransferase gene from
RT   Habanero Chile.";
RL   (In) Plant Gene Register PGR98-170.
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF081214; AAC78475.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
SQ   SEQUENCE   359 AA;  39636 MW;  4A836904EF6D7119 CRC64;

  Query Match             19.9%;  Score 375.5;  DB 1;  Length 359;
  Best Local Similarity   29.3%;  Pred. No. 6e-23;
  Matches   98;  Conservative   67;  Mismatches  124;  Indels   45;  Gaps    9;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIH 93
              ||  ||||: :|:     |  |: ||| : |   |||  :  ::    : | :: :
Db         32 LKATVELDLLEIMAKSGPGAFISPSELAAQL---PTKNPEAPVML-----DRMFRLLATY 83

Qy         94 DNIEA------------YALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----L 135
                 :             |:|    : | |: : :|:||::            ||    :
Db         84 SVLNCTLRTLPDGRVERLYSLAPVCKFLTKNGDGVSIAPIL----------LMNQDKVLM 133

Qy        136 KRWVHEEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWV 190
              : | |  |  :      |  : |   :::   || :|| ||  |:  : :      :
Db        134 ESWYHLTDAVLDGGVPFNKAYGMTTFEYHGTDPRFNKVFNCGMSDHTTLSMKKILEDYTG 193
```

```
Qy        191 FEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIP 250
              |||| |||||||||| |  :|   :| :|  :  : |:|: :      :  ||||||  :|
Db        194 FEGLNSIVDVGGGTGATVNMIVSKYPSIKGINFDLPHVIRDAPSYPGVEQVGGDMFVSVP 253

Qy        251 KADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKL 310
              ||||: :| : |:|:|: |:|:|:|| ||:      |||:::: :: |  |    |: |:
Db        254 KADAIFMKWICHDWSDDHCIKLLKNCYEALPA---NGKVIIVECILPEAPDTSAATKSKV 310

Qy        311 LMDVHM-ACIINGKERKEEDWKKLFMEAGFQSYK 343
                |: | |    |||| |:|:: |     |  ::
Db        311 HGDIIMLAHNPGGKERTEKDFEALANWGWFSRFR 344


RESULT 12
COM1_OCIBA
ID   COM1_OCIBA     STANDARD;      PRT;   361 AA.
AC   Q9XGW0;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase 1 (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase 1) (COMT-1) (CAOMT-1).
GN   COMT1.
OS   Ocimum basilicum (Sweet basil).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Lamiales; Lamiaceae; Nepetoideae; Ocimeae;
OC   Ocimum.
OX   NCBI_TaxID=39350;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. EXM-1;
RA   Wang J., Dudareva N., Kish C.M., Simon J.E., Lewinsohn E.,
RA   Pichersky E.;
RT   "Nucleotide sequences of two cDNAs encoding caffeic acid O-
RT   methyltransferases from sweet basil (Ocimum basilicum).";
RL   (In) Plant Gene Register PGR99-105.
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF154917; AAD38189.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
SQ   SEQUENCE   361 AA;  39528 MW;  D6ABC3DC837AAC9E CRC64;

  Query Match             19.7%;  Score 371.5;  DB 1;  Length 361;
  Best Local Similarity   29.7%;  Pred. No. 1.3e-22;
  Matches   97;  Conservative   69;  Mismatches  128;  Indels   33;  Gaps   11;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR---QVQSLMRYLAHNGFFE-- 88
              ||  :|||: ::|     |  ::  :| :  |:| |       :  ::| |       |
Db         34 LKSAIELDLLELIKKSGAGAFVSPVDLAA--QLPTTNPDAHVMLDRILRLLTSYAILECR 91

Qy         89 IVRIHD-NIE-AYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVHE 141
                :  : |  :|  | |    : | |: : :|:||:             ||    :: | |
Db         92 LKTLPDGGVERLYGLAPVCKFLTKNEDGVSMAPLT----------LMNQDKVLMESWYHL 141

Qy        142 EDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLES 196
                |  |     |  : |   :::   || :|| ||: |:  | :      :    |:||::
Db        142 SDAVVDGGIPFNKAYGMTAFEYHGTDPRFNKVFNQGMSNHSTITMKKILETYTGFDGLKT 201

Qy        197 IVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVL 256
              :|||||||| |  :|   :| :|  :  : |:|:|:      :  ||||||  :|| ||:
Db        202 VVDVGGGTGATLNMIVSKYPSIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPKGDAIF 261

Qy        257 LKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM 316
              :| : |:|:|  |:| |:|| :|:    : |||:: : |: |  |    |:  : :|| |
Db        262 MKWICHDWSDEHCVKFLKNCYDAL---PQNGKVILAECVLPEAPDTGLATKNVVHIDVIM 318

Qy        317 -ACIINGKERKEEDWKKLFMEAGFQSY 342
               |    |||| |:::: |   |||: :
Db        319 LAHNPGGKERTEKEFQGLAKAAGFKQF 345


RESULT 13
COM1_POPTM
ID   COM1_POPTM     STANDARD;      PRT;   365 AA.
AC   Q00763; Q43094;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase 1 (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase 1) (COMT-1) (CAOMT-1).
GN   OMT1.
OS   Populus tremuloides (Quaking aspen).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Malpighiales; Salicaceae; Populus.
OX   NCBI_TaxID=3693;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 165-184; 335-346 AND 349-359.
RC   TISSUE=Xylem;
RX   MEDLINE=92032785; PubMed=1932694;
RA   Bugos R.C., Chiang V.L.C., Campbell W.H.;
RT   "cDNA cloning, sequence analysis and seasonal expression of lignin-
RT   bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase of
RT   aspen.";
RL   Plant Mol. Biol. 17:1203-1215(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95288373; PubMed=7770534;
RA   Tsai C.-J., Podila G.K., Chiang V.L.C.;
RT   "Nucleotide sequence of a Populus tremuloides gene encoding
RT   bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase.";
RL   Plant Physiol. 107:1459-1459(1995).
RN   [3]
RP   REVISIONS.
RA   Tsai C.-J., Mielke M.R., Podila G.K., Chiang V.L.C.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- TISSUE SPECIFICITY: Xylem.
CC   -!- PTM: The N-terminus is blocked.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; X62096; CAA44006.1; -.
DR   EMBL; U13171; AAB61731.1; -.
DR   PIR; S18568; S18568.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
SQ   SEQUENCE   365 AA;  39805 MW;  A6CECDEA4E0007CD CRC64;

  Query Match             19.7%;  Score 371.5;  DB 1;  Length 365;
  Best Local Similarity   29.7%;  Pred. No. 1.3e-22;
  Matches   97;  Conservative   69;  Mismatches  132;  Indels   29;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFF-- 87
              ||  :|||: :|:      |  :: ||: | |   |||       :  ::| ||
Db         36 LKTAIELDLLEIMAKAGPGAFLSTSEIASHL---PTKNPDAPVMLDRILRLLASYSILTC 92

Qy         88 --------EIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRW 138
                    :: |:      | |   : | |: : :|::|:     :     :|  ||
Db         93 SLKDLPDGKVERL------YGLAPVCKFLTKNEDGVSVSPLCLMNQDKVLMESWYYLKDA 146

Qy        139 VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
              : :  :  |  : |   :::   || :|| ||: |:  | :      :    |||| |:|
Db        147 ILDGGIP-FNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYKGFEGLTSLV 205

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
              |||||||     |   :| :| :  : |:|:|:      :  |||||  :|||||| :|
Db        206 DVGGGTGAVVNTIVSKYPSIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPKADAVFMK 265

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
              : |:|:|  |:| |:|| :|:    : |||:::: ::    |    |:  : :|| | |
Db        266 WICHDWSDAHCLKFLKNCYDAL---PENGKVILVECILPVAPDTSLATKGVVHVDVIMLA 322

Qy        318 CIINGKERKEEDWKKLFMEAGFQSYKI 344
                 |||| |:::: |   |||| :::
Db        323 HNPGGKERTEKEFEGLAKGAGFQGFEV 349


RESULT 14
COM1_POPKI
ID   COM1_POPKI     STANDARD;      PRT;   365 AA.
AC   Q43046;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase 1 (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase 1) (COMT-1) (CAOMT-1).
GN   HOMT1.
OS   Populus kitakamiensis (Aspen).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Malpighiales; Salicaceae; Populus.
OX   NCBI_TaxID=3698;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hayakawa T., Nanto K., Kawai S., Katayama Y., Morohoshi N.;
RT   "Molecular cloning and tissue-specific expression of two genes that
RT   encode caffeic acid O-methyltransferases from Populus kitakamiensis.";
RL   Plant Sci. 113:157-165(1996).
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; D49710; BAA08558.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
SQ   SEQUENCE   365 AA;  39791 MW;  D6005B10FE55B83C CRC64;

  Query Match             19.5%;  Score 368.5;  DB 1;  Length 365;
  Best Local Similarity   29.7%;  Pred. No. 2.2e-22;
  Matches   97;  Conservative   68;  Mismatches  133;  Indels   29;  Gaps    8;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGFF-- 87
              ||  :|||: :|:      |  :: ||: | |   |||       :  ::| ||
Db         36 LKTAIELDLLEIMAKAGPGAFLSTSEIASHL---PTKNPDAPVMLDRILRLLASYSILTC 92

Qy         88 --------EIVRIHDNIEAYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQLKRW 138
                    :: |:      | |   : | |: : :|::|:     :     :|  ||
Db         93 SLKDLPDGKVERL------YGLAPVCKFLTKNEDGVSVSPLCLMNQDKVLMESWYYLKDA 146

Qy        139 VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLESIV 198
              : |  :  |  : |   :::   || :|| ||: |:  | :      :    |||| |:|
Db        147 ILEGGIP-FNKAYGMTAFEYHGTDPRFNKVFNKGMSDHSTITMKKILETYKGFEGLTSLV 205

Qy        199 DVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAVLLK 258
              |||||||     |   :| :| :  : |:|:|:      :  |||||  :|||||| :|
Db        206 DVGGGTGAVVNTIVSKYPSIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSVPNADAVFMK 265

Qy        259 LVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVHM-A 317
              : |:|:|  |:| |:|| :|:    : |||:::: ::    |    |:  : :|| | |
Db        266 WICHDWSDAHCLKFLKNCYDAL---PENGKVILVECILPVAPDTSLATKGVVHVDVIMLA 322

Qy        318 CIINGKERKEEDWKKLFMEAGFQSYKI 344
                 |||| |:::: |   |||| :::
Db        323 HNPGGKERTEKEFEGLANGAGFQGFEV 349


RESULT 15
COMT_PRUDU
ID   COMT_PRUDU     STANDARD;      PRT;   365 AA.
AC   Q43609;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caffeic acid 3-O-methyltransferase (EC 2.1.1.68) (S-adenosysl-L-
DE   methionine:caffeic acid 3-O-methyltransferase) (COMT) (CAOMT).
GN   COMT1.
OS   Prunus dulcis (Almond) (Prunus amygdalus).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX   NCBI_TaxID=3755;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Texas; TISSUE=Root;
RA   Garcia-Mas J., Messeguer R., Arus P., Puigdomenech P.;
RT   "The caffeic acid O-methyltransferase from Prunus amygdalus.";
RL   (In) Plant Gene Register PGR95-006.
CC   -!- FUNCTION: Catalyzes the conversion of caffeic acid to ferulic acid
CC       and of 5-hydroxyferulic acid to sinapic acid. The resulting
CC       products may subsequently be converted to the corresponding
CC       alcohols that are incorporated into lignins.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3,4-dihydroxy-trans-
CC       cinnamate = S-adenosyl-L-homocysteine + 3-methoxy-4-hydroxy-trans-
```

```
CC       cinnamate.
CC   -!- PATHWAY: Lignin biosynthesis.
CC   -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
CC       2. COMT subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X83217; CAA58218.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Lignin biosynthesis; Transferase; Methyltransferase.
SQ   SEQUENCE   365 AA;  39762 MW;  C8C4BFEE7E0180E7 CRC64;

  Query Match           19.5%;  Score 367.5;  DB 1;  Length 365;
  Best Local Similarity   30.7%;  Pred. No. 2.7e-22;
  Matches  104;  Conservative   65;  Mismatches  133;  Indels   37;  Gaps   13;

Qy         34 LKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR----QVQSLMRYLAHNGF--F 87
              ||  :|||: :|:     |  :: ::: | |   |||      :  ::| ||     :
Db         36 LKAAIELDLLEIMAKAGPGVFLSPTDIASQL---PTKNPDAPVMLDRMLRLLASYSILTY 92

Qy         88 EIVRIHD-NIE-AYALTAASELLVKSSE-LSLAPMVEYFLEPNCQGAWNQ----LKRWVH 140
               :  : |  :|  | |    : | |: | :|:||:        |    ||    |: | |
Db         93 SLRTLADGKVERLYGLGPVCKFLTKNEEGVSIAPL--------C--LMNQDKVLLESWYH 142

Qy        141 EEDLTV-----FEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAFRDCNWVFEGLE 195
              :|  :     |  : |   :::   || :|| ||  ||  | :      :    ||||
Db        143 LKDAVLEGGIPFNKAYGMTAFEYHGTDPRFNKVFNRGMADHSTITMKKILETYKGFEGLT 202

Qy        196 SIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPKADAV 255
              |:|||||||    :|   :| :| :  : |:|:|:      :  ||||||  :|| ||:
Db        203 SVVDVGGGTGAVLNMIVSKYPSIKGINFDLPHVIEDAPQYPGVEHVGGDMFVSVPKGDAI 262

Qy        256 LLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKLLMDVH 315
              :| : |:|:|  |:| |:||  |:      |||:: : ::   |    |:  : :||
Db        263 FMKWICHDWSDEHCLKFLKNCYAAL---PDNGKVILGECILPVAPDSSLATKGVVHIDVI 319

Qy        316 M-ACIINGKERKEEDWKKLFMEAGFQSYKI--SPFTGYL 351
              | |   |||| |:::: |   |||| : :  | |  |:
Db        320 MLAHNPGGKERTEQEFQALAKGAGFQGFNVACSAFNTYV 358


Search completed: November 10, 2003, 03:19:19
Job time : 25 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 10, 2003, 03:03:30 ; Search time 107 Seconds
(without alignments)
863.391 Million cell updates/sec

Title: US-09-868-547-4
Perfect score: 1886
Sequence: 1 MASSLNNGRKASEIFQGQAL.....FQSYKISPFTGYLSLIEIYP 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1295	68.7	352	10	O22309	O22309 medicago sa
2	1289	68.3	352	10	O22308	O22308 medicago sa
3	1285	68.1	352	10	O24529	O24529 medicago sa
4	1237.5	65.6	343	10	O22555	O22555 medicago sa
5	867.5	46.0	360	10	O24305	O24305 pisum sativ
6	863	45.8	367	10	Q8L5K8	Q8l5k8 rosa hybrid
7	863	45.8	367	10	Q8GU24	Q8gu24 rosa chinen
8	847.5	44.9	366	10	Q8L5K7	Q8l5k7 rosa hybrid
9	847.5	44.9	366	10	Q8GU23	Q8gu23 rosa chinen
10	837	44.4	354	10	O24249	O24249 prunus arme
11	821	43.5	348	10	Q8GU21	Q8gu21 rosa hybrid
12	817	43.3	348	10	Q8GU22	Q8gu22 rosa hybrid
13	757	40.1	356	10	O65859	O65859 prunus dulc
14	734	38.9	357	10	Q93WU2	Q93wu2 ocimum basi
15	726.5	38.5	347	10	Q8GSN0	Q8gsn0 catharanthu
16	723.5	38.4	356	10	Q93WU3	Q93wu3 ocimum basi
17	716	38.0	348	10	Q8GSN1	Q8gsn1 catharanthu
18	692.5	36.7	383	10	Q9ST37	Q9st37 pyrus pyrif
19	652.5	34.6	390	10	Q43771	Q43771 hordeum vul
20	608	32.2	366	10	Q8VWJ6	Q8vwj6 oryza sativ
21	603	32.0	366	10	Q8VWG4	Q8vwg4 oryza sativ
22	588	31.2	378	10	Q8RV92	Q8rv92 oryza sativ
23	579	30.7	373	10	Q9ZRC1	Q9zrc1 triticum ae
24	578	30.6	360	10	Q8LI64	Q8li64 oryza sativ
25	538	28.5	381	10	Q43096	Q43096 pinus taeda
26	535.5	28.4	382	10	Q9SYU3	Q9syu3 pinus radia
27	529.5	28.1	382	10	O24287	O24287 pinus radia
28	527	27.9	382	10	Q9T003	Q9t003 arabidopsis
29	517	27.4	351	10	Q8H9A8	Q8h9a8 coptis japo
30	485	25.7	325	10	Q9T002	Q9t002 arabidopsis
31	431.5	22.9	341	10	Q8LI58	Q8li58 oryza sativ
32	426	22.6	363	10	O49010	O49010 zea mays (m
33	398.5	21.1	364	10	Q9SYR8	Q9syr8 thalictrum
34	398.5	21.1	364	10	Q9XE91	Q9xe91 thalictrum
35	392.5	20.8	362	10	Q9XE92	Q9xe92 thalictrum
36	391.5	20.8	359	10	Q9FQY8	Q9fqy8 capsicum an
37	390.5	20.7	350	10	Q8LL87	Q8ll87 coffea cane
38	390.5	20.7	364	10	Q42958	Q42958 nicotiana t
39	387.5	20.5	367	10	Q96424	Q96424 glycyrrhiza
40	386.5	20.5	350	10	Q8LL88	Q8ll88 coffea cane
41	386.5	20.5	364	10	Q04065	Q04065 nicotiana t
42	384	20.4	372	10	P93324	P93324 medicago sa
43	383.5	20.3	359	10	P93088	P93088 capsicum an
44	383.5	20.3	363	10	Q8W013	Q8w013 catharanthu
45	379.5	20.1	362	10	Q9XE90	Q9xe90 thalictrum

ALIGNMENTS

RESULT 1
O22309
ID O22309 PRELIMINARY; PRT; 352 AA.
AC O22309;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 7-O-methyltransferase.
GN 7-IOMT(9).
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98145455; PubMed=9484461;
RA He X.Z., Reddy J.T., Dixon R.A.;
RT "Stress responses in alfalfa (Medicago sativa L). XXII. cDNA cloning
RT and characterization of an elicitor-inducible isoflavone 7-O-
RT methyltransferase.";
RL Plant Mol. Biol. 36:43-54(1998).
DR EMBL; AF000976; AAC49927.1; -.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR001077; O_Metransf2.
DR Pfam; PF00891; Methyltransf_2; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 352 AA; 39540 MW; 955683CC90B94ACD CRC64;

Query Match 68.7%; Score 1295; DB 10; Length 352;
Best Local Similarity 68.2%; Pred. No. 9.5e-102;
Matches 244; Conservative 55; Mismatches 53; Indels 6; Gaps 4;

```
Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              ||||: |||| ||||: |||||||: |||| ||| | ::||:|| |:||:||: | |
Db          1 MASSI-NGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVGMNIPNII--HNHGKPISLSNL 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120
              ||||||| :|   |: |||||||||||||:   :  |:|||| ||||||: |:| |||||
```

```
Db         58 VSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEE--ESYALTVASELLVRGSDLCLAPMV 115

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180

Db        116 ECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMASDSKLI 175

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240

Db        176 NLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTY 235

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300

Db        236 VGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVTNDGKRGKVTIIDMVINEKK 295

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358

Db        296 DENQVTQIKLLMDVNMAC-LNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLIEIYP 352
```

RESULT 2
O22308
ID O22308 PRELIMINARY; PRT; 352 AA.
AC O22308;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 7-O-methyltransferase.
GN 7-IOMT(6).
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98145455; PubMed=9484461;
RA He X.Z., Reddy J.T., Dixon R.A.;
RT "Stress responses in alfalfa (Medicago sativa L). XXII. cDNA cloning
RT and characterization of an elicitor-inducible isoflavone 7-O-
RT methyltransferase.";
RL Plant Mol. Biol. 36:43-54(1998).
DR EMBL; AF000975; AAC49926.1; -.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR001077; O_Metransf2.
DR Pfam; PF00891; Methyltransf_2; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 352 AA; 39604 MW; 31B95226966C1C98 CRC64;

Query Match 68.3%; Score 1289; DB 10; Length 352;
Best Local Similarity 67.9%; Pred. No. 3.1e-101;
Matches 243; Conservative 56; Mismatches 53; Indels 6; Gaps 4;

```
Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60

Db          1 MASSI-NGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNII--QNHGKPISLSNL 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120

Db         58 VSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEE--ESYALTVASELLVRGSDLCLAPMV 115

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180

Db        116 ECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMASDSKLI 175

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240

Db        176 NLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTY 235

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300

Db        236 VGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVTNDGKRGKVTIIDMVIDEKK 295

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358

Db        296 DENQVTQIKLLMDVNMAC-LNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLIEIYP 352
```

RESULT 3
O24529
ID O24529 PRELIMINARY; PRT; 352 AA.
AC O24529;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Isoflavone-O-methytransferase.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98145455; PubMed=9484461;
RA He X.Z., Reddy J.T., Dixon R.A.;
RT "Stress responses in alfalfa (Medicago sativa L). XXII. cDNA cloning
RT and characterization of an elicitor-inducible isoflavone 7-O-
RT methyltransferase.";
RL Plant Mol. Biol. 36:43-54(1998).
DR EMBL; U97125; AAC49928.1; -.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR001077; O_Metransf2.
DR Pfam; PF00891; Methyltransf_2; 1.
KW Transferase.
SQ SEQUENCE 352 AA; 39603 MW; 31B95228986C1296 CRC64;

Query Match 68.1%; Score 1285; DB 10; Length 352;
Best Local Similarity 67.6%; Pred. No. 6.7e-101;
Matches 242; Conservative 57; Mismatches 53; Indels 6; Gaps 4;

```
Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60

Db          1 MASSI-NGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNII--QNHGKPISLSNL 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120

Db         58 VSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEE--ESYALTVASELLVRGSDLCLAPMV 115

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180

Db        116 ECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMASDSKLI 175

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240

Db        176 NLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTY 235

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300

Db        236 VGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVTNDGKRGKVTIIDMVIDKKK 295

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358

Db        296 DENQVTQIKLLMDVNMAC-LNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLIEIYP 352
```

RESULT 4
O22555
ID O22555 PRELIMINARY; PRT; 343 AA.
AC O22555;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE O-methytransferase.
GN IOMT2003.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX   NCBI_TaxID=3879;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   He X.Z., Dixon R.A.;
RT   "A cDNA of an additional member of the isoflavone-O-methyltransferase
RT   (IOMT2) gene family in Medicago sativa (Accession No. AF023481)
RT   (PGR97-170).";
RL   Plant Physiol. 115:1289-1289(1997).
DR   EMBL; AF023481; AAB88294.1; -.
DR   InterPro; IPR001077; O_Metransf2.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Transferase.
SQ   SEQUENCE   343 AA;  38433 MW;  296665211415C7BB CRC64;

  Query Match             65.6%;  Score 1237.5;  DB 10;  Length 343;
  Best Local Similarity   66.5%;  Pred. No. 7e-97;
  Matches  238;  Conservative   55;  Mismatches   50;  Indels   15;  Gaps    5;

Qy          1 MASSLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSEL 60
              ||||: |||| |||||: ||||||||:  |||| ||| ||::||:||   :||:||: | |
Db          1 MASSI-NGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNII--QNHGKPISLSNL 57

Qy         61 VSILQVPPTKTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMV 120
              |||||||  :|   |: ||||||||||||||:   :  |:|||| ||||||: |:| |||||
Db         58 VSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEE--ESYALTVASELLVRGSDLCLAPMV 115

Qy        121 EYFLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML 180
              |  |:|   |::::||:|::|||||:| |:|:: ||||::|:| || |||:||| ||::
Db        116 ECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMASDSKLI 175

Qy        181 NLAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTF 240
              ||| |||::||:||||||||||||| ||||||| ||||||:| :|| ||||||||||||:
Db        176 NLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTY 235

Qy        241 VGGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENK 300
              ||||||  || |||||||         ||::||: |||:: : | ||| :|| ||:| |
Db        236 VGGDMFTSIPNADAVLLK---------DCLRILKKCKEAVTNDGKRGKVTIIDMVIDEKK 286

Qy        301 DERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              || |||::||||||:||| :||||| ||:|||||:||||| ||||| ||:||||||||
Db        287 DENQVTQIKLLMDVNMAC-LNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLIEIYP 343


RESULT 5
O24305
ID   O24305      PRELIMINARY;      PRT;   360 AA.
AC   O24305;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   6a-hydroxymaackiain methyltransferase.
GN   HMM6.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Alaska;
RX   MEDLINE=98009990; PubMed=9349277;
RA   Wu Q., Preisig C.L., VanEtten H.D.;
RT   "Isolation of the cDNAs encoding (+)6a-hydroxymaackiain 3-O-
RT   methyltransferase, the terminal step for the synthesis of the
RT   phytoalexin pisatin in Pisum sativum.";
RL   Plant Mol. Biol. 35:551-560(1997).
DR   EMBL; U69554; AAC49856.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Methyltransferase; Transferase.
SQ   SEQUENCE   360 AA;  40351 MW;  97882D1189A93888 CRC64;

  Query Match             46.0%;  Score 867.5;  DB 10;  Length 360;
  Best Local Similarity   48.7%;  Pred. No. 2e-65;
  Matches  174;  Conservative   60;  Mismatches  116;  Indels    7;  Gaps    3;

Qy          7 NGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQV 66
              || : ||::  |  ||||:  |: |  ||  :|| | | |  |:||:|:|  || | |::
Db          6 NGSEESELYHAQIHLYKHVYNFVSSMALKSAMELGIADAI--HNHGKPMTLPELSSSLKL 63

Qy         67 PPTKTRQVQSLMRYLAHNGFFEIVRIHDN----IEAYALTAASELLVKSSELSLAPMVEY 122
                |:|   :   :| | |||||    :  |      || || :|:|||      |: :|:
Db         64 HPSKVNILYRFLRLLTHNGFFAKTTVKSNEGEEETAYVLTPSSKLLVSGKSTCLSSLVKG 123

Qy        123 FLEPNCQGAWNQLKRWVHEE-DLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLN 181
              | |:    |   |:| ||: : |:|| : |  :|||:|||      | :||| ||::
Db        124 ALHPSSLDMWGVSKKWFHEDKEQTLFECATGENYWDFLNKDSDSLSMFQDAMAADSRLFK 183

Qy        182 LAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFV 241
              || ::   |||||||:||| |||  ||:| |||| :|| | ::| || ||:|: || ||
Db        184 LAIQENKHVFEGLESLVDVAGGTGGVAKLIHEAFPHIKCTVFDQPQVVGNLTGNENLNFV 243

Qy        242 GGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD 301
              |||||| :| ||||||| |||:|||   :|||:| ||||| : | |||::||  |:|| |
Db        244 GGDMFKSVPSADAVLLKWVLHDWNDELSLKILKNSKEAISHKGKDGKVIIIDISIDENSD 303

Qy        302 ERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :| :|||:|  || |  :  |||| :::|:||  :|||  |||:|  |: ||||:||
Db        304 DRGLTELQLEYDVVMLTMFLGKERTKKEWEKLIYDAGFSRYKITPICGFKSLIEVYP 360


RESULT 6
Q8L5K8
ID   Q8L5K8      PRELIMINARY;      PRT;   367 AA.
AC   Q8L5K8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orcinol O-methyltransferase.
GN   OOMT1.
OS   Rosa hybrid cultivar.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX   NCBI_TaxID=128735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=fc0496;
RX   MEDLINE=22167322; PubMed=12177504;
RA   Lavid N., Wang J., Shalit M., Guterman I., Bar E., Beuerle T.,
RA   Menda N., Shafir S., Zamir D., Adam Z., Vainstein A., Weiss D.,
RA   Pichersky E., Lewinsohn E.;
RT   "O-methyltransferases involved in the biosynthesis of volatile
RT   phenolic derivatives in rose petals.";
RL   Plant Physiol. 129:1899-1907(2002).
DR   EMBL; AF502433; AAM23004.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   367 AA;  41278 MW;  9739F3E2D09FEF7C CRC64;

  Query Match             45.8%;  Score 863;  DB 10;  Length 367;
  Best Local Similarity   46.2%;  Pred. No. 5e-65;
  Matches  164;  Conservative   68;  Mismatches  119;  Indels    4;  Gaps    3;

Qy          6 NNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQ 65
              :||  ::|:   || :: |:  ||:|  ||  ::| |||||  : || |:| ||| | |
Db         15 SNGEHSNELLHAQAHIWNHIFSFINSMSLKSAIQLGIPDII--NKHGYPMTLSELTSALP 72

Qy         66 VPPTKTRQVQSLMRYLAHNGFFEIVRI-HDNIEAYALTAASELLVKSSELSLAPMVEYFL 124
              : |||:  |  ||| | |:|||   ::   : | | || ||:||:|   ||| | :   |
Db         73 IHPTKSHSVYRLMRILVHSGFFAKKKLSKTDEEGYTLTDASQLLLKDHPLSLTPYLTAML 132
```

```
Qy        125 EPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML-NLA 183
              :|      || |  |   :| | |: : |  |||: |  |:    ||:||| |:::: ::
Db        133 DPVLTNPWNYLSTWFQNDDPTPFDTAHGMTFWDYGNHQPSIAHLFNDAMASDARLVTSVI 192

Qy        184 FRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGG 243
                ||  ||||||:||||||||  || | :||| ::| ||: |:|| :| || || : ||
Db        193 INDCKGVFEGLESLVDVGGGTGTLAKAIADAFPHIECTVLDLPHVVADLQGSKNLKYTGG 252

Qy        244 DMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDER 303
              |||: :| || |||| :||:|:| :|:|||| : ||:|: | |||::|| ::   | :
Db        253 DMFEAVPPADTVLLKWILHDWSDEECIKILERSRVAITGKEKKGKVIIIDMMMENQKGDE 312

Qy        304 QVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :  | :|  |: |  :: |||| |::| ||| :||| |||:| :|  ||||:||
Db        313 ESIETQLFFDMLMMALVGGKERNEKEWAKLFTDAGFSDYKITPISGLRSLIEVYP 367


RESULT 7
Q8GU24
ID   Q8GU24      PRELIMINARY;      PRT;   367 AA.
AC   Q8GU24;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orcinol O-methyltransferase.
GN   OOMT1.
OS   Rosa chinensis.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX   NCBI_TaxID=74649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Old Blush;
RA   Cock J.Mark, Scalliet G., Hugueney P.;
RT   "Characterisation of a novel O-methyltransferase involved in the
RT   biosynthesis of 3,5-dimethoxytoluene and 1,3,5-trimethoxybenzene, two
RT   major scent components of rose flowers.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ439741; CAD29458.1; -.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   367 AA;  41278 MW;  9739F3E2D09FEF7C CRC64;

  Query Match             45.8%;  Score 863;  DB 10;  Length 367;
  Best Local Similarity   46.2%;  Pred. No. 5e-65;
  Matches  164;  Conservative   68;  Mismatches  119;  Indels    4;  Gaps    3;

Qy          6 NNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQ 65
              :||  ::|:   || :: |: .||:|  ||  ::| ||||  : || |:| ||| | |
Db         15 SNGEHSNELLHAQAHIWNHIFSFINSMSLKSAIQLGIPDII--NKHGYPMTLSELTSALP 72

Qy         66 VPPTKTRQVQSLMRYLAHNGFFEIVRI-HDNIEAYALTAASELLVKSSELSLAPMVEYFL 124
              : |||:  |  ||| | |:|||   ::   : | | || ||:||:|   ||| | :   |
Db         73 IHPTKSHSVYRLMRILVHSGFFAKKKLSKTDEEGYTLTDASQLLLKDHPLSLTPYLTAML 132

Qy        125 EPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML-NLA 183
              :|      || |  |   :| | |: : |  |||: |  |:    ||:||| |:::: ::
Db        133 DPVLTNPWNYLSTWFQNDDPTPFDTAHGMTFWDYGNHQPSIAHLFNDAMASDARLVTSVI 192

Qy        184 FRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGG 243
                ||  ||||||:||||||||  || | :||| ::| ||: |:|| :| || || : ||
Db        193 INDCKGVFEGLESLVDVGGGTGTLAKAIADAFPHIECTVLDLPHVVADLQGSKNLKYTGG 252

Qy        244 DMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDER 303
              |||: :| || |||| :||:|:| :|:|||| : ||:|: | |||::|| ::   | :
Db        253 DMFEAVPPADTVLLKWILHDWSDEECIKILERSRVAITGKEKKGKVIIIDMMMENQKGDE 312

Qy        304 QVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :  | :|  |: |  :: |||| |::| ||| :||| |||:| :|  ||||:||
Db        313 ESIETQLFFDMLMMALVGGKERNEKEWAKLFTDAGFSDYKITPISGLRSLIEVYP 367


RESULT 8
Q8L5K7
ID   Q8L5K7      PRELIMINARY;      PRT;   366 AA.
AC   Q8L5K7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orcinol O-methyltransferase.
GN   OOMT2.
OS   Rosa hybrid cultivar.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX   NCBI_TaxID=128735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=fc0895;
RX   MEDLINE=22167322; PubMed=12177504;
RA   Lavid N., Wang J., Shalit M., Guterman I., Bar E., Beuerle T.,
RA   Menda N., Shafir S., Zamir D., Adam Z., Vainstein A., Weiss D.,
RA   Pichersky E., Lewinsohn E.;
RT   "O-methyltransferases involved in the biosynthesis of volatile
RT   phenolic derivatives in rose petals.";
RL   Plant Physiol. 129:1899-1907(2002).
DR   EMBL; AF502434; AAM23005.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   366 AA;  41226 MW;  CAF4E17D3230D29E CRC64;

  Query Match             44.9%;  Score 847.5;  DB 10;  Length 366;
  Best Local Similarity   45.6%;  Pred. No. 1e-63;
  Matches  162;  Conservative   67;  Mismatches  121;  Indels    5;  Gaps    3;

Qy          6 NNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQ 65
              :||  ::|:   || :: |:  ||:|  ||  ::| |||||: |    |:| ||| | |
Db         15 SNGEHSNELLHAQAHIWNHIFSFINSMSLKSAIQLGIPDIINKHG---PMTLSELTSALP 71

Qy         66 VPPTKTRQVQSLMRYLAHNGFFEIVRI-HDNIEAYALTAASELLVKSSELSLAPMVEYFL 124
              : |||:  |  ||| | |:|||   ::   : | | || ||:||:|   ||| | :   |
Db         72 IHPTKSHSVYRLMRILVHSGFFAKKKLSKTDEEGYTLTDASQLLLKDHPLSLTPFLTAML 131

Qy        125 EPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML-NLA 183
              :|      || |  |   || | |: : |  |||: |  |:    ||:||| |:::: ::
Db        132 DPVLTTPWNYLSTWFQNEDPTPFDTAHGMTFWDYGNHQPSIAHLFNDAMASDARLVTSVI 191

Qy        184 FRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGG 243
                ||  ||||||:||||||||  || | :||| ::| ||: |:|| :| || || : ||
Db        192 IDDCKGVFEGLESLVDVGGGTGTVAKAIADAFPHIECTVLDLPHVVADLQGSKNLKYTGG 251

Qy        244 DMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDER 303
              |||: :| || |||| :||:||| :|:|||:  : ||: : | |||::|| ::   | :
Db        252 DMFEAVPPADTVLLKWILHDWNDEECIKILKRSRVAITSKDKKGKVIIIDMMMENQKGDE 311

Qy        304 QVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :  | :|  |: |  :: |:|| |::| ||| :||| |||:|  |  ||||:||
Db        312 ESIETQLFFDMLMMALVRGQERNEKEWAKLFTDAGFSDYKITPILGLRSLIEVYP 366


RESULT 9
Q8GU23
ID   Q8GU23      PRELIMINARY;      PRT;   366 AA.
AC   Q8GU23;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orcinol O-methyltransferase.
GN   OOMT2.
OS   Rosa chinensis.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

```
OC   eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX   NCBI_TaxID=74649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Old Blush;
RA   Cock J.Mark, Scalliet G., Hugueney P.;
RT   "Characterisation of a novel O-methyltransferase involved in the
RT   biosynthesis of 3,5-dimethoxytoluene and 1,3,5-trimethoxybenzene, two
RT   major scent components of rose flowers.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ439742; CAD29459.1; -.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   366 AA;  41226 MW;  CAF4E17D3230D29E CRC64;

  Query Match             44.9%;  Score 847.5;  DB 10;  Length 366;
  Best Local Similarity   45.6%;  Pred. No. 1e-63;
  Matches  162;  Conservative   67;  Mismatches  121;  Indels    5;  Gaps    3;

Qy          6 NNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQ 65
              :||  ::|:   || :: |:  ||:|  ||  ::| |||||: |    |:| ||| | |
Db         15 SNGEHSNELLHAQAHIWNHIFSFINSMSLKSAIQLGIPDIINKHG---PMTLSELTSALP 71

Qy         66 VPPTKTRQVQSLMRYLAHNGFFEIVRI-HDNIEAYALTAASELLVKSSELSLAPMVEYFL 124
                : |||:  |  ||| | |:|||   ::   : | | || ||:||:|   ||| | :   |
Db         72 IHPTKSHSVYRLMRILVHSGFFAKKKLSKTDEEGYTLTDASQLLLKDHPLSLTPFLTAML 131

Qy        125 EPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML-NLA 183
              :|     || |  |   || | |: : |  |||:  |  |:    ||:||| |:::: ::
Db        132 DPVLTTPWNYLSTWFQNEDPTPFDTAHGMTFWDYGNHQPSIAHLFNDAMASDARLVTSVI 191

Qy        184 FRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGG 243
                ||  |||||||:||||||||  || | :||| ::| ||: |:|| :| || || : ||
Db        192 IDDCKGVFEGLESLVDVGGGTGTVAKAIADAFPHIECTVLDLPHVVADLQGSKNLKYTGG 251

Qy        244 DMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDER 303
              |||: :| || |||| :||:||| :|:|||:  : ||: : | |||::|| ::   | :
Db        252 DMFEAVPPADTVLLKWILHDWNDEECIKILKRSRVAITSKDKKGKVIIIDMMMENQKGDE 311

Qy        304 QVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :  | :|  |: |  :: |:|| |::| ||| :|||   |||:|  |  ||||:||
Db        312 ESIETQLFFDMLMMALVRGQERNEKEWAKLFTDAGFSDYKITPILGLRSLIEVYP 366


RESULT 10
O24249
ID   O24249      PRELIMINARY;      PRT;   354 AA.
AC   O24249;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Methyltransferase.
OS   Prunus armeniaca (Apricot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX   NCBI_TaxID=36596;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bergeron; TISSUE=Mesocarp, and Exocarp;
RA   Mbeguie-A-Mbeguie D., Gomez R.-M., Fils-Lycaon B.;
RT   "Sequence of an O-Methyltransferase from Apricot Fruit (Accession No.
RT   U82011). Gene Expression During Fruit Ripening (PGR97-118).";
RL   Plant Physiol. 114:1569-1569(1997).
DR   EMBL; U82011; AAB71213.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Transferase.
SQ   SEQUENCE   354 AA;  39384 MW;  7F6350F76F196CDB CRC64;

  Query Match             44.4%;  Score 837;  DB 10;  Length 354;
  Best Local Similarity   45.2%;  Pred. No. 7.7e-63;
  Matches  159;  Conservative   66;  Mismatches  123;  Indels    4;  Gaps    3;

Qy          9 RKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPP 68
              | : |: | || :: |:  ||:|  ||  |:|||||:|    ||||:| ||||| | : |
Db          5 RASHELLQAQAHIWNHIFSFINSLSLKCAVQLDIPDVI--QKHGQPMTLSELVSALPISP 62

Qy         69 TKTRQVQSLMRYLAHNGFFEIVRIHD-NIEAYALTAASELLVKSSELSLAPMVEYFLEPN 127
              ||   :  ||| | |:|||    :     : | || || ||:|: : :| . | :   | |
Db         63 TKAHFIPRLMRILVHSGFFAKESLSGCGEQGYILTDASALLLKDNPMSARPFLLAMLSPI 122

Qy        128 CQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML-NLAFRD 186
                :  |  |   :: | | |  |   ||::|:||     ||:||| |:::: :|   |
Db        123 LTDPYQYLTTWFQNDNPTPFHVVNGMTCWDYVNQDPTLAHFFNDAMASDAQLISSLVIDD 182

Qy        187 CNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMF 246
              |  ||:|::|:|||||||  || | :||| :|| ||: |:|| :| || || :| ||||
Db        183 CKEVFQGVDSLVDVGGGTGTVAKSIADAFPHMKCTVLDLPHVVADLKGSKNLEYVAGDMF 242

Qy        247 KCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVT 306
              : :| |||: || :||:|:|  :|:|||| || |::  | | |||:::  :     :::
Db        243 EAVPAADAIFLKWILHDWSDEECVKILERCKAAVTREGKKGKVIIVEMTVENKNTDKESG 302

Qy        307 ELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              | :|  |:||  :  |||| |::| ||| :|||  |||:|  |: ||||:||
Db        303 ETQLFFDMHMMVMSTGKERNEKEWAKLFSDAGFSQYKITPLFGFKSLIEVYP 354


RESULT 11
Q8GU21
ID   Q8GU21      PRELIMINARY;      PRT;   348 AA.
AC   Q8GU21;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orcinol O-methyltransferase (Fragment).
GN   OOMT4.
OS   Rosa hybrid cultivar.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX   NCBI_TaxID=128735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Lady Hillington;
RA   Cock J.Mark, Scalliet G., Hugueney P.;
RT   "Characterisation of a novel O-methyltransferase involved in the
RT   biosynthesis of 3,5-dimethoxytoluene and 1,3,5-trimethoxybenzene, two
RT   major scent components of rose flowers.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ439744; CAD29556.1; -.
KW   Transferase; Methyltransferase.
FT   NON_TER      1      1
FT   NON_TER    348    348
SQ   SEQUENCE   348 AA;  38982 MW;  24391E70DDB33C09 CRC64;

  Query Match             43.5%;  Score 821;  DB 10;  Length 348;
  Best Local Similarity   45.4%;  Pred. No. 1.7e-61;
  Matches  158;  Conservative   67;  Mismatches  115;  Indels    8;  Gaps    4;

Qy          6 NNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQ 65
              :||  ::|:   || :: |:  ||:|  ||  ::| |||||  : || |:| ||| | |
Db          5 SNGEHSNELLHAQAHIWNHIFSFINSMSLKSAIQLGIPDII--NKHGYPMTLSELSSALP 62

Qy         66 VPPTKTRQVQSLMRYLAHNGFF---EIVRIHDNIEAYALTAASELLVKSSELSLAPMVEY 122
              : |||:  |  ||| | |:|||   :: :| :  | | || ||:|:|:|   ||| | :
Db         63 IHPTKSHSVYRLMRILVHSGFFAKKKLSKIDE--EGYTLTDASQLLLKDHPLSLTPFLTA 120

Qy        123 FLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML-N 181
                |:|     || |  |   :| | |: : |  |||: |  |     ||:||| |:::: :
Db        121 MLDPVLTKPWNYLSTWFQNDDPTPFDTTHGMTFWDYGNHQPNIAHLFNDAMASDARLVTS 180

Qy        182 LAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFV 241
              :  ||  |||||||:||||||||  || | :||| ::| ||: |:|| :| || || :
```

```
Db        181 VIIDDCKGVFEGLESLVDVGGGTGTVAKAIADAFPHIECTVLDLPHVVGDLQGSKNLKYT 240

Qy        242 GGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD 301
              |||||: :| || |||| :||:||| :|:|||:  : ||: : | |||::|| ::   |
Db        241 GGDMFEAVPPADTVLLKWILHDWNDEECIKILKRSRVAITSKDKKGKVIIIDMMMENQKG 300

Qy        302 ERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTG 349
              : :  | :| :|: |  :: |||| |::| ||| :|||  |||:|  |
Db        301 DEESIETQLFLDMLMMALVGGKERNEKEWAKLFTDAGFSDYKITPILG 348


RESULT 12
Q8GU22
ID   Q8GU22      PRELIMINARY;      PRT;   348 AA.
AC   Q8GU22;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orcinol O-methyltransferase (Fragment).
GN   OOMT3.
OS   Rosa hybrid cultivar.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX   NCBI_TaxID=128735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Lady Hillington;
RA   Cock J.Mark, Scalliet G., Hugueney P.;
RT   "Characterisation of a novel O-methyltransferase involved in the
RT   biosynthesis of 3,5-dimethoxytoluene and 1,3,5-trimethoxybenzene, two
RT   major scent components of rose flowers.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ439743; CAD29555.1; -.
KW   Transferase; Methyltransferase.
FT   NON_TER       1      1
FT   NON_TER     348    348
SQ   SEQUENCE   348 AA;  39034 MW;  0AFD4D325A0DBB1D CRC64;

  Query Match             43.3%;  Score 817;  DB 10;  Length 348;
  Best Local Similarity   45.1%;  Pred. No. 3.8e-61;
  Matches  156;  Conservative   67;  Mismatches  119;  Indels    4;  Gaps    3;

Qy          6 NNGRKASEIPQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQ 65
              :||  ::|:    || :: |:  ||:|  ||  ::| |||||  : :| |:| ||| | |
Db          5 SNGEHSNELLHAQAHIWNHIFSFINSMSLKSAIQLGIPDII--NKYGYPMTLSELTSALP 62

Qy         66 VPPTKTRQVQSLMRYLAHNGFFEIVRI-HDNIEAYALTAASELLVKSSELSLAPMVEYFL 124
              : |||:  |  ||| | |:|||   ::   : | | || ||:||:|   ||| | :   |
Db         63 IHPTKSHSVYRLMRILVHSGFFAKKKLSKTDEEGYTLTDASQLLLKDHPLSLTPYLTAML 122

Qy        125 EPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQML-NLA 183
              :|     || |  |   :| | |: : |  |||: |  |:    ||:||| |:::: ::
Db        123 DPVLTNPWNYLSTWFQNDDPTPFDTAHGMTFWDYGNHQPSIAHLFNDAMASDARLVTSVI 182

Qy        184 FRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGG 243
                ||  |||||||:||||||||  || | :||| ::| ||: |:|| :| || || : ||
Db        183 IDDCKGVFEGLESLVDVGGGTGTMAKAIADAFPHIECTVLDLPHVVADLQGSKNLKYTGG 242

Qy        244 DMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDER 303
              |||: :| || |||| :||:||| :|:|||:  : ||: : | |||::|| ::   | :
Db        243 DMFEAVPPADTVLLKWILHDWNDQECIKILKRSRVAITSKDKKGKVIIIDMMMENQKGDE 302

Qy        304 QVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTG 349
              :  | :|  |: |  :: |||| |::| ||| :|||  |||:|  |
Db        303 ESIETQLFFDMLMMALVGGKERNEKEWAKLFTDAGFSDYKITPILG 348


RESULT 13
O65859
ID   O65859      PRELIMINARY;      PRT;   356 AA.
AC   O65859;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   O-methyltransferase.
OS   Prunus dulcis (Almond) (Prunus amygdalus).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX   NCBI_TaxID=3755;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Texas;
RA   Suelves M., Puigdomenech P.;
RT   "Specific mRNA accumulation of a gene coding for an O-
RT   methyltransferase in almond (Prunus amygdalus, Batsch) flower
RT   tissues.";
RL   Plant Sci. 134:79-88(1998).
DR   EMBL; AJ223151; CAA11131.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Methyltransferase; Transferase.
SQ   SEQUENCE   356 AA;  40206 MW;  4CC6F169513F4EA9 CRC64;

  Query Match             40.1%;  Score 757;  DB 10;  Length 356;
  Best Local Similarity   42.3%;  Pred. No. 4.8e-56;
  Matches  151;  Conservative   70;  Mismatches  130;  Indels    6;  Gaps    4;

Qy          5 LNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSIL 64
              |:|   :: : | || ::  :  ||:   ||  |:| |||||     || |:: |:|:| |
Db          3 LSNEMSSANLLQAQAHIWNCIFSFINPMSLKCAVQLGIPDII--KKHGNPMSLSDLISAL 60

Qy         65 QVPPTKTRQVQSLMRYLAHNGFFEIVRIH--DNIEAYALTAASELLVKSSELSLAPMVEY 122
              : | |:  |  ||| | |:|||   ::   |  | | || || ||:|   ||  | :
Db         61 PIHPKKSNCVYRLMRILVHSGFFCRQKLSELDEEEGYVLTDASRLLLKDDPLSARPFLLG 120

Qy        123 FLEPNCQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLN- 181
              |:|      |:    |   :| |    : || ||||   :|: :  ||:||| |:::::
Db        121 ALDPFMTKPWHYFSTWFQNDDPTACVTAHGTTFWDFGCLEPSLSHIFNDAMASDARLISK 180

Qy        182 LAFRDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFV 241
              :   :  |||||||:|||||| |   | | : || ::|:| : |:|| :| || || |
Db        181 VVSNEYKGVFEGLESLVDVGGGIGTMPKAIADVFPHVECIVFDLPHVVADLKGSENLKFF 240

Qy        242 GGDMFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKD 301
              |||||: || ||:|:| :||:|:| :|:|||:  ||||: : | |||:::| |: :  |
Db        241 GGDMFEAIPPTDAILMKWILHDWSDEECIKILQRSKEAITRKEKKGKVIIVDMVMKQKGD 300

Qy        302 ERQVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
              :: : | :|  |: |  :  |||| |::| ||| :|||  |||:|  |   |||:||
Db        301 DQSI-ETQLFFDMLMMVLFTGKERTEKEWAKLFSDAGFSDYKITPICGLRYLIEVYP 356


RESULT 14
Q93WU2
ID   Q93WU2      PRELIMINARY;      PRT;   357 AA.
AC   Q93WU2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Eugenol O-methyltransferase.
OS   Ocimum basilicum (sweet basil).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Ocimum.
OX   NCBI_TaxID=39350;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. EMX-1; TISSUE=Peltate glandular trichome;
RA   Gang D.R., Lavid N., Zubieta C., Chen F., Beuerle T., Lewinsohn E.,
RA   Noel J.P., Pichersky E.;
RT   "Characterization of Phenylpropene O-Methyltransferases from Sweet
RT   Basil: Facile Change of Substrate Specificity and Convergent Evolution
RT   within a Plant OMT Family.";
```

```
RL   Plant Cell 0:0-0(2002).
DR   EMBL; AF435008; AAL30424.1; -.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR001077; O_Metransf2.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00891; Methyltransf_2; 1.
KW   Methyltransferase; Transferase.
SQ   SEQUENCE   357 AA;  40236 MW;  FC050C48BF6D719C CRC64;

  Query Match             38.9%;  Score 734;  DB 10;  Length 357;
  Best Local Similarity   41.5%;  Pred. No. 4.4e-54;
  Matches  145;  Conservative   75;  Mismatches  121;  Indels    8;  Gaps    5;

Qy         13 EIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPTKTR 72
Db         14 QLLQAQVHVWNHMYAFANSMSLKCAIQLGIPDIL--HKHGRPMTLSQLLQSIPINKEKTQ 71

Qy         73 QVQSLMRYLAHNGFFEIVRIHDNIE-AYALTAASELLVKSSELSLAPMVEYFLEPNCQGA 131
Db         72 CFQRLMRALVNSNFFIEENNSNNQEVCYWLTPASCLLLKEAPLTVTPLVQVVLDPTFTNP 131

Qy        132 WNQLKRW-VHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLNLAF-RDCNW 189
Db        132 WHHMSEWFTHEKHATQFEAANGCTFWEKLANEPSKGRFFDEAMSCDSRLIAHVFTKDYKH 191

Qy        190 VFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCI 249
Db        192 VIEGIRTLVDVGGGNGTMAKAIVEAMPTIKCTVIDLPHVVAGLESTDNLNYIGGDMFQSI 251

Qy        250 PKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELK 309
Db        252 PSADAILLKSIIHDWDDVEGLKILKKCKDAV---VMGGKVIIIDVVVGVNHDIDEVLEDQ 308

Qy        310 LLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
Db        309 LHFDMAMMCYFNAKERTMSEWEKLIYDAGFKSYKLTPAFGVRSLIEAYP 357


RESULT 15
Q8GSN0
ID   Q8GSN0      PRELIMINARY;      PRT;   347 AA.
AC   Q8GSN0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   O-methyltransferase.
OS   Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae;
OC   Catharanthus.
OX   NCBI_TaxID=4058;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22371605; PubMed=12482447;
RA   Cacace S., Schroeder G., Wehinger E., Strack D., Schmid J.,
RA   Schroeder J.;
RT   "A flavonol O-methyltransferase from Catharanthus roseus performing
RT   two sequential methylations.";
RL   Phytochemistry 62:127-137(2003).
DR   EMBL; AY127569; AAM97498.1; -.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   347 AA;  38789 MW;  7FB24AC0BC97AB35 CRC64;

  Query Match             38.5%;  Score 726.5;  DB 10;  Length 347;
  Best Local Similarity   41.7%;  Pred. No. 1.8e-53;
  Matches  148;  Conservative   73;  Mismatches  117;  Indels   17;  Gaps    8;

Qy         10 KASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSHSHGQPITFSELVSILQVPPT 69
Db          4 ETAEIRKAQAQYSRHVFSFVSTSCLKCAIQLEIPEAI--HNHGKPMTLSDLTNSLPINPS 61

Qy         70 KTRQVQSLMRYLAHNGFFEIVRIHDNIEAYALTAASELLVKSSELSLAPMVEYFLEPN-- 127
Db         62 KAPYIHRLMRILVAAGYFS----EEPKNVYSLTSLSRILVKNQPLNLR---EFVLSANEI 114

Qy        128 --CQGAWNQLKRWVHEEDLTVFEVSLGTPFWDFINKDPAYNKSFNEAMACDSQMLN-LAF 184
Db        115 AEVEG-WNALSEWFQNDVATAFQTAHGKTYWEYLSQD-KYGKNFDQLMATDSLLISKLLI 172

Qy        185 RDCNWVFEGLESIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGD 244
Db        173 PDYNYLFEGLISLVDVGGGTGTLAGAVAKAFPNLKCTVFEQPHVIADLEAKGNLEFVGGD 232

Qy        245 MFKCIPKADAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVI-NENKDER 303
Db        233 MFEKIPSANAILLKSVLHDWKDEDSVKILKNCKKAIPEKEKGGKVIVIDIVLMDSKKHDN 292

Qy        304 QVTELKLLMDVHMACIINGKERKEEDWKKLFMEAGFQSYKISPFTGYLSLIEIYP 358
Db        293 PLVKSQISGDMDMMVSMGAKERTEEEWAALFKEAGFSGYKIFPMLDFRSPIEVYP 347


Search completed: November 10, 2003, 03:21:16
Job time : 110 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 9, 2003, 23:30:19 ; Search time 4620 Seconds
(without alignments)
10271.681 Million cell updates/sec

Title: US-09-868-547-3
Perfect score: 1160
Sequence: 1 gtttgctattatggcttctt.....ataaaagttatctttgtgtt 1160

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	697.8	60.2	1121	8	AB091685	AB091685 Glycyrrhi
2	647	55.8	1284	8	MSAF000975	AF000975 Medicago
3	645.4	55.6	1231	6	AX050436	AX050436 Sequence
4	645.4	55.6	1231	8	MSU97125	U97125 Medicago sa
5	642.2	55.4	1227	8	MSAF000976	AF000976 Medicago
6	594.2	51.2	1194	8	AF023481	AF023481 Medicago
7	337.6	29.1	1349	8	AB091684	AB091684 Glycyrrhi
8	321.6	27.7	1337	8	PSU69554	U69554 Pisum sativ
c 9	318.8	27.5	111587	2	AC136140	AC136140 Medicago
10	318	27.4	1363	8	AB091686	AB091686 Lotus jap
11	305.6	26.3	1257	8	PAU82011	U82011 Prunus arme
12	302.8	26.1	1253	8	AF502434	AF502434 Rosa hybr
13	302.8	26.1	1265	8	RCH439742	AJ439742 Rosa chin
14	302.4	26.1	1252	8	AF502433	AF502433 Rosa hybr
15	302.4	26.1	1275	8	RCH439741	AJ439741 Rosa chin
16	287.6	24.8	1043	8	RHY439744	AJ439744 Rosa hybr
17	286	24.7	1043	8	RHY439743	AJ439743 Rosa hybr
18	274.2	23.6	1218	8	PAAJ3151	AJ223151 Prunus am
19	209.2	18.0	1041	6	E28578	E28578 Norcoclauri
20	209.2	18.0	1267	8	D29811	D29811 Coptis japo
21	206	17.8	1071	8	AF435007	AF435007 Ocimum ba
22	205.6	17.7	1074	8	AF435008	AF435008 Ocimum ba
23	195	16.8	83007	8	AC123575	AC123575 Medicago
24	180.4	15.6	1280	8	D29812	D29812 Coptis japo
25	179.8	15.5	112530	2	AC140068	AC140068 Medicago
26	178.4	15.4	1050	6	E27546	E27546 (s)-3'-Hydr
27	152.2	13.1	1402	8	HVOMET	X77467 H.vulgare L
28	147.4	12.7	1354	8	AB014456	AB014456 Pyrus pyr
29	146.8	12.7	1659	8	AY127568	AY127568 Catharant
30	146.6	12.6	1279	8	PTU39301	U39301 Pinus taeda
31	144.8	12.5	1703	8	AY127569	AY127569 Catharant
32	142	12.2	1268	8	MZEOMT	L14063 Zea mays O-
33	131.6	11.3	1378	8	AB073908	AB073908 Coptis ja
34	131	11.3	1421	8	PRU70873	U70873 Pinus radia
35	130	11.2	1396	8	AF239740	AF239740 Vitis vin
36	125.8	10.8	1503	8	AF212316	AF212316 Capsicum
37	124.2	10.7	1377	8	CAU83789	U83789 Capsicum an
38	122.2	10.5	951	6	AX660124	AX660124 Sequence
39	121	10.4	1469	8	AF064694	AF064694 Thalictru
40	119.4	10.3	1306	8	AF064695	AF064695 Thalictru
41	119.2	10.3	1316	8	AF064693	AF064693 Thalictru
42	119.2	10.3	1360	8	AF064696	AF064696 Thalictru
43	119.2	10.3	1458	8	TAU76384	U76384 Triticum ae
44	118	10.2	1098	6	AX259373	AX259373 Sequence
45	118	10.2	1317	8	ALFSALMCAT	M63853 Medicago sa

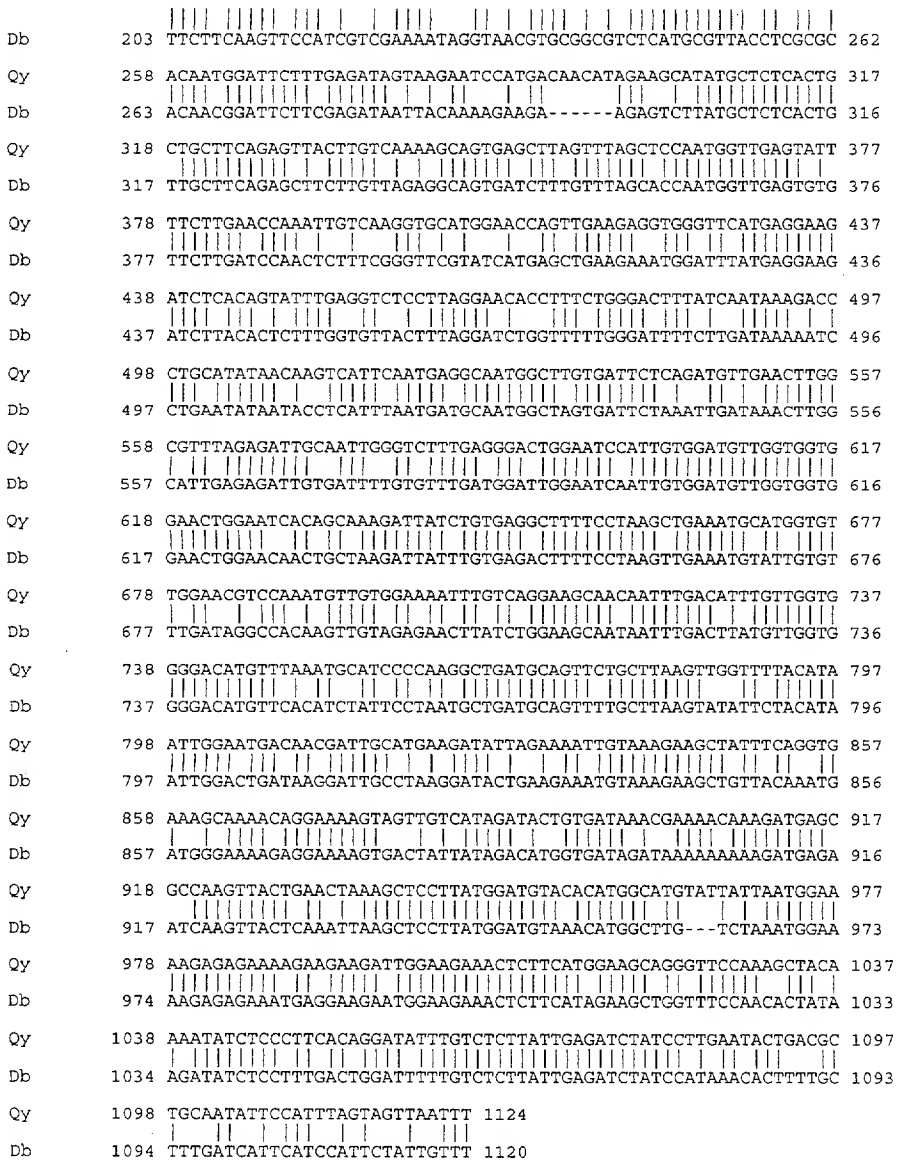
ALIGNMENTS

RESULT 1
AB091685
LOCUS AB091685 1121 bp mRNA linear PLN 04-MAR-2003
DEFINITION Glycyrrhiza echinata D7OMT mRNA for S-adenosyl-L-methionine: daidzein 7-0-methyltransferase, complete cds.
ACCESSION AB091685
VERSION AB091685.1 GI:28804593
KEYWORDS .
SOURCE Glycyrrhiza echinata
ORGANISM Glycyrrhiza echinata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrrhiza.
REFERENCE 1

```
  AUTHORS   Akashi,T., Sawada,Y., Shimada,N., Sakurai,N., Aoki,T. and Ayabe,S.
  TITLE     cDNA cloning and biochemical characterization of
            s-adenosyl-L-methionine: 2,7,4'-trihydroxyisoflavanone
            4'-o-methyltransferase, a critical enzyme of the legume
            isoflavonoid phytoalexin pathway
  JOURNAL   Plant Cell Physiol. 44 (2), 103-112 (2003)
  MEDLINE   22497945
   PUBMED   12610212
REFERENCE   2  (bases 1 to 1121)
  AUTHORS   Akashi,T., Aoki,T. and Ayabe,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-SEP-2002) Shin-ichi Ayabe, Nihon University,
            Department of Applied Biological Sciences; Kameino 1866, Fujisawa,
            Kanagawa 252-8510, Japan (E-mail:ayabe@brs.nihon-u.ac.jp,
            Tel:81-466-84-3703, Fax:81-466-80-1141)
FEATURES             Location/Qualifiers
     source          1. .1121
                     /organism="Glycyrrhiza echinata"
                     /mol_type="mRNA"
                     /db_xref="taxon:46348"
                     /cell_line="Ak-1"
                     /note="cloning vector: lambda ZapII~cultured cells"
     gene            1. .1121
                     /gene="D7OMT"
     CDS             48. .1121
                     /gene="D7OMT"
                     /codon_start=1
                     /product="S-adenosyl-L-methionine: daidzein
                     7-0-methyltransferase"
                     /protein_id="BAC58012.1"
                     /db_xref="GI:28804594"
                     /translation="MASSINGRKPSEIFQGQALLYRHIYAFIDSMCLKWIVELDIPNI
                     IHNHGKPITVSELVSILKVPQTKAGNVQRIMRYMAHNGFFERVRIQEEQEENEAYALT
                     AASELLVKGSELCLAPMVECVLDPTLSGSYHQLKKWIYEEDLTLFGVSLGSHFWEFLN
                     ENPEYNKSFNDAMASDSQMINLALRDCNSGFEGVESIVDVGGGIGTTAKIICDTFPNL
                     KCIVFDRPKVVENLSGTNNLSYVGGDMFQSVPKADAVLLKWILHNWTDNDCRRILEKC
                     KEAVSSDGEKGKVIIIEMVINENQDEHEITGTKLLMDVNMACLNGKERSEEEWKKLFI
                     EAGFRDYKISPLTGFLSLIEVYP"
BASE COUNT      339 a    210 c    267 g    305 t
ORIGIN

  Query Match             60.2%;  Score 697.8;  DB 8;  Length 1121;
  Best Local Similarity   80.1%;  Pred. No. 3.2e-148;
  Matches  864;  Conservative    0;  Mismatches  197;  Indels   18;  Gaps    3;

Qy         18 CTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACA 77
Db         52 CTTCTTCAATCAATGGGCGCAAACCAAGTGAGATCTTTCAAGGCCAGGCTCTCTTGTACA 111

Qy         78 AACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATAC 137
Db        112 GACACATATACGCCTTCATAGATTCCATGTGTCTGAAATGGATCGTTGAGCTAGACATAC 171

Qy        138 CCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAA 197
Db        172 CAAACATAATCCAC------AACCATGGAAAACCCATTACTGTTTCAGAGTTGGTGTCAA 225

Qy        198 TTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCAC 257
Db        226 TCCTAAAAGTTCCACAAACCAAAGCTGGGAACGTGCAGCGTATCATGCGCTACATGGCAC 285

Qy        258 ACAATGGATTCTTTGAGATAGTAAGAATC---------CATGACAACATAGAAGCATATG 308
Db        286 ACAATGGTTTCTTTGAGAGAGTGAGAATCCAAGAAGAGCAAGAAGAGAACGAAGCATATG 345

Qy        309 CTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGG 368
Db        346 CTCTAACAGCTGCTTCAGAGCTACTTGTGAAAGGCTCTGAGCTCTGTTTGGCTCCAATGG 405

Qy        369 TTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTC 428
Db        406 TTGAGTGTGTTCTTGACCCAACTCTCTCAGGTTCATACCATCAGCTCAAGAAGTGGATCT 465

Qy        429 ATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCA 488
Db        466 ATGAGGAAGATCTCACACTCTTTGGTGTCTCTTTGGGATCACATTTCTGGGAATTCCTCA 525

Qy        489 ATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGT 548
Db        526 ACGAGAACCCTGAATATAACAAGTCCTTCAATGATGCAATGGCCAGTGATTCTCAGATGA 585

Qy        549 TGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATG 608
Db        586 TAAACTTGGCTTTGAGAGATTGCAATTCGGGGTTTGAGGGAGTGGAATCCATTGTGGATG 645

Qy        609 TTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAAT 668
Db        646 TTGGTGGTGGAATTGGAACCACAGCCAAGATTATCTGTGACACATTTCCCAACTTGAAAT 705

Qy        669 GCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACAT 728
Db        706 GCATAGTCTTTGACCGTCCAAAGGTTGTGGAGAATTTGTCTGGAACCAACAATTTGAGTT 765

Qy        729 TTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGG 788
Db        766 ATGTTGGTGGGGACATGTTCCAATCTGTTCCCAAGGCTGATGCTGTTCTACTTAAGTGGA 825

Qy        789 TTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTA 848
Db        826 TTTTGCATAATTGGACAGATAACGATTGCAGAAGGATATTGGAAAAATGTAAAGAAGCAG 885

Qy        849 TTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACA 908
Db        886 TTTCAAGTGATGGGGAAAAGGGAAAAGTAATTATTATAGAAATGGTGATAAATGAAAATC 945

Qy        909 AAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTA 968
Db        946 AGGATGAGCATGAAATCACTGGAACAAAGCTCCTTATGGATGTAAACATGGCATG---TC 1002

Qy        969 TTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCC 1028
Db       1003 TTAATGGAAAAGAGAGAAGTGAGGAAGAATGGAAGAAGCTCTTCATTGAAGCAGGCTTCC 1062

Qy       1029 AAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGA 1087
Db       1063 GAGACTACAAAATATCTCCTTTGACTGGATTTTTGTCCCTTATTGAGGTCTATCCTTAA 1121
```

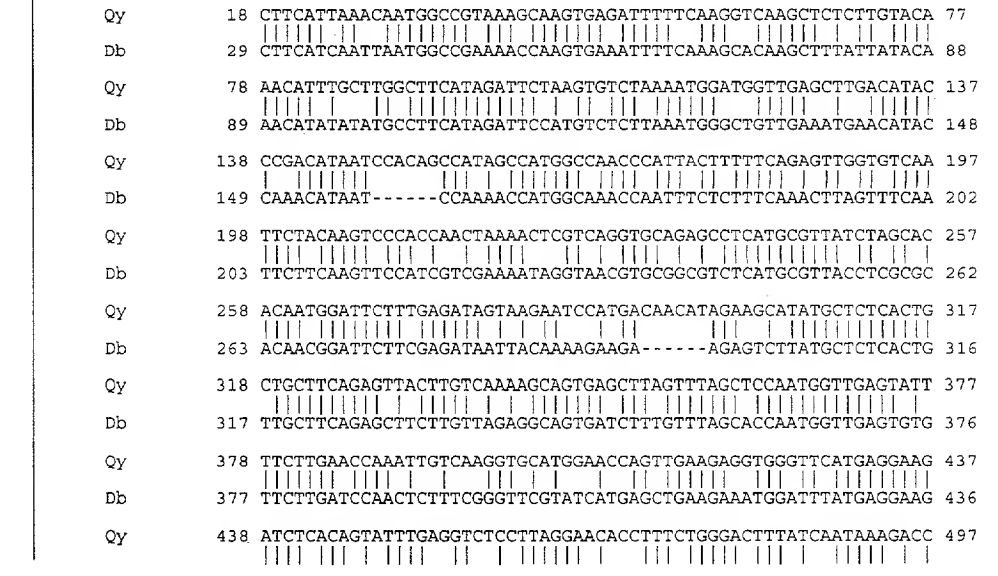
```
RESULT 2
MSAF000975
LOCUS       MSAF000975              1284 bp    mRNA    linear   PLN 25-FEB-1998
DEFINITION  Medicago sativa 7-O-methyltransferase (7-IOMT(6)) mRNA, complete
            cds.
ACCESSION   AF000975
VERSION     AF000975.1  GI:2580581
KEYWORDS    .
SOURCE      Medicago sativa
  ORGANISM  Medicago sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 1284)
  AUTHORS   He,X.Z., Reddy,J.T. and Dixon,R.A.
  TITLE     Stress responses in alfalfa (Medicago sativa L). XXII. cDNA cloning
            and characterization of an elicitor-inducible isoflavone
            7-O-methyltransferase
  JOURNAL   Plant Mol. Biol. 36 (1), 43-54 (1998)
  MEDLINE   98145455
   PUBMED   9484461
REFERENCE   2  (bases 1 to 1284)
  AUTHORS   He,X.Z., Reddy,J.T., Guo,J.Z. and Dixon,R.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-1997) Plant Biology Divsion, The Samuel Roberts
            Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
FEATURES             Location/Qualifiers
```

```
     source          1. .1284
                     /organism="Medicago sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:3879"
     gene            1. .1284
                     /gene="7-IOMT(6)"
     CDS             30. .1088
                     /gene="7-IOMT(6)"
                     /codon_start=1
                     /product="7-O-methyltransferase"
                     /protein_id="AAC49926.1"
                     /db_xref="GI:2580582"
                     /translation="MASSINGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNI
                     IQNHGKPISLSNLVSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEEESYALTVASEL
                     LVRGSDLCLAPMVECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEY
                     NTSFNDAMASDSKLINLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVF
                     DRPQVVVENLSGSNNLTYVGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVT
                     NDGKRGKVTIIDMVIDEKKDENQVTQIKLLMDVNMACLNGKERNEEEWKKLFIEAGFQ
                     HYKISPLTGFLSLIEIYP"
BASE COUNT      421 a    189 c    244 g    430 t
ORIGIN

  Query Match           55.8%;  Score 647;  DB 8;  Length 1284;
  Best Local Similarity  76.1%;  Pred. No. 1.1e-136;
  Matches  842;  Conservative    0;  Mismatches  250;  Indels   15;  Gaps    3;

Qy         18 CTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACA 77
Db         34 CTTCATCAATTAATGGCCGAAAACCAAGTGAAATTTTCAAAGCACAAGCTTTATTATACA 93

Qy         78 AACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATAC 137
Db         94 AACATATATATGCCTTCATAGATTCCATGTCTCTTAAATGGGCTGTTGAAATGAACATAC 153

Qy        138 CCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAA 197
Db        154 CAAACATAAT------CCAAAACCATGGCAAACCAATTTCTCTTTCAAACTTAGTTTCAA 207

Qy        198 TTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCAC 257
Db        208 TTCTTCAAGTTCCATCGTCGAAAATAGGTAACGTGCGGCGTCTCATGCGTTACCTCGCGC 267

Qy        258 ACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTG 317
Db        268 ACAACGGATTCTTCGAGATAATTACAAAAGAAGA------AGAGTCTTATGCTCTCACTG 321

Qy        318 CTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATT 377
Db        322 TTGCTTCAGAGCTTCTTGTTAGAGGCAGTGATCTTTGTTTAGCACCAATGGTTGAGTGTG 381

Qy        378 TTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAG 437
Db        382 TTCTTGATCCAACTCTTTCGGGTTCGTATCATGAGCTGAAGAAATGGATTTATGAGGAAG 441

Qy        438 ATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACC 497
Db        442 ATCTTACACTCTTTGGTGTTACTTTAGGATCTGGTTTTTGGGATTTTCTTGATAAAAATC 501

Qy        498 CTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGG 557
Db        502 CTGAATATAATACCTCATTTAATGATGCAATGGCTAGTGATTCTAAATTGATAAACTTGG 561

Qy        558 CGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTG 617
Db        562 CATTGAGAGATTGTGATTTTGTGTTTGATGGATTGGAATCAATTGTGGATGTTGGTGGTG 621

Qy        618 GAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGT 677
Db        622 GAACTGGAACAACTGCTAAGATTATTTGTGAGACTTTTCCTAAGTTGAAATGTATTGTGT 681

Qy        678 TGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTG 737
Db        682 TTGATAGGCCACAAGTTGTAGAGAACTTATCTGGAAGCAATAATTTGACTTATGTTGGTG 741

Qy        738 GGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATA 797
Db        742 GGGACATGTTCACATCTATTCCTAATGCTGATGCAGTTTTGCTTAAGTATATTCTACATA 801

Qy        798 ATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTG 857
Db        802 ATTGGACTGATAAGGATTGCCTAAGGATACTGAAGAAATGTAAAGAAGCTGTTACAAATG 861

Qy        858 AAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGC 917
Db        862 ATGGGAAAAGAGGAAAAGTGACTATTATAGACATGGTGATAGATGAAAAAAAAGATGAGA 921

Qy        918 GCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAA 977
Db        922 ATCAAGTTACTCAAATTAAGCTCCTTATGGATGTAAACATGGCTTG---TCTAAATGGAA 978

Qy        978 AAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACA 1037
Db        979 AAGAGAGAAATGAGGAAGAATGGAAGAAACTCTTCATAGAAGCTGGTTTCCAACACTATA 1038

Qy       1038 AAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGC 1097
Db       1039 AGATATCTCCTTTGACTGGATTTTTGTCTCTTATTGAGATCTATCCATAAACACTTTTGC 1098

Qy       1098 TGCAATATTCCATTTAGTAGTTAATTT 1124
Db       1099 TTTGATCATTCATCCATTCTATTGTTT 1125


RESULT 3
AX050436
LOCUS       AX050436               1231 bp    DNA     linear   PAT 12-JAN-2001
DEFINITION  Sequence 1 from Patent WO0071736.
ACCESSION   AX050436
VERSION     AX050436.1  GI:12226643
KEYWORDS    .
SOURCE      Medicago sativa
  ORGANISM  Medicago sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1
  AUTHORS   Dixon,R.A. and He,X.Z.
  TITLE     Isoflavonoid methylation enzyme
  JOURNAL   Patent: WO 0071736-A 1 30-NOV-2000;
            The Samuel Roberts Noble Foundation, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1231
                     /organism="Medicago sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3879"
BASE COUNT      402 a    183 c    235 g    411 t
ORIGIN

  Query Match           55.6%;  Score 645.4;  DB 6;  Length 1231;
  Best Local Similarity  76.0%;  Pred. No. 2.6e-136;
  Matches  841;  Conservative    0;  Mismatches  251;  Indels   15;  Gaps    3;

Qy         18 CTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACA 77
Db         29 CTTCATCAATTAATGGCCGAAAACCAAGTGAAATTTTCAAAGCACAAGCTTTATTATACA 88

Qy         78 AACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATAC 137
Db         89 AACATATATATGCCTTCATAGATTCCATGTCTCTTAAATGGGCTGTTGAAATGAACATAC 148

Qy        138 CCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAA 197
Db        149 CAAACATAAT------CCAAAACCATGGCAAACCAATTTCTCTTTCAAACTTAGTTTCAA 202

Qy        198 TTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCAC 257
```



RESULT 4
MSU97125
LOCUS MSU97125 1231 bp mRNA linear PLN 25-FEB-1998
DEFINITION Medicago sativa isoflavone-O-methytransferase mRNA, complete cds.
ACCESSION U97125
VERSION U97125.1 GI:2580539
KEYWORDS .
SOURCE Medicago sativa
ORGANISM Medicago sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 1231)
AUTHORS He,X.Z., Reddy,J.T. and Dixon,R.A.
TITLE Stress responses in alfalfa (Medicago sativa L). XXII. cDNA cloning and characterization of an elicitor-inducible isoflavone 7-O-methyltransferase
JOURNAL Plant Mol. Biol. 36 (1), 43-54 (1998)
MEDLINE 98145455
PUBMED 9484461
REFERENCE 2 (bases 1 to 1231)
AUTHORS He,X., Reddy,J.T. and Dixon,R.A.
TITLE Direct Submission
JOURNAL Submitted (11-APR-1997) Plant Biology Division, The Samuel Roberts Noble Foundation Inc., 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
FEATURES Location/Qualifiers
source 1. .1231
/organism="Medicago sativa"
/mol_type="mRNA"
/db_xref="taxon:3879"
CDS 25. .1083
/note="7-IOMT; 7-O-methyltransferase"
/codon_start=1
/product="isoflavone-O-methytransferase"
/protein_id="AAC49928.1"
/db_xref="GI:2580540"
/translation="MASSINGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNI
IQNHGKPISLSNLVSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEEESYALTVASEL
LVRGSDLCLAPMVECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEY
NTSFNDAMASDSKLINLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVF
DRPQVVENLSGSNNLTYVGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVT
NDGKRGKVTIIDMVIDKKKDENQVTQIKLLMDVNMACLNGKERNEEEWKKLFIEAGFQ
HYKISPLTGFLSLIEIYP"
BASE COUNT 402 a 183 c 235 g 411 t
ORIGIN

Query Match 55.6%; Score 645.4; DB 8; Length 1231;
Best Local Similarity 76.0%; Pred. No. 2.6e-136;
Matches 841; Conservative 0; Mismatches 251; Indels 15; Gaps 3;



```
Db        437 ATCTTACACTCTTTGGTGTTACTTTAGGATCTGGTTTTTGGGATTTTCTTGATAAAAATC 496

Qy        498 CTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGG 557

Db        497 CTGAATATAATACCTCATTTAATGATGCAATGGCTAGTGATTCTAAATTGATAAACTTGG 556

Qy        558 CGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTG 617

Db        557 CATTGAGAGATTGTGATTTTGTGTTTGATGGATTGGAATCAATTGTGGATGTTGGTGGTG 616

Qy        618 GAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGT 677

Db        617 GAACTGGAACAACTGCTAAGATTATTTGTGAGACTTTTCCTAAGTTGAAATGTATTGTGT 676

Qy        678 TGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTG 737

Db        677 TTGATAGGCCACAAGTTGTAGAGAACTTATCTGGAAGCAATAATTTGACTTATGTTGGTG 736

Qy        738 GGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATA 797

Db        737 GGGACATGTTCACATCTATTCCTAATGCTGATGCAGTTTTGCTTAAGTATATTCTACATA 796

Qy        798 ATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTG 857

Db        797 ATTGGACTGATAAGGATTGCCTAAGGATACTGAAGAAATGTAAAGAAGCTGTTACAAATG 856

Qy        858 AAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGC 917

Db        857 ATGGGAAAAGAGGAAAAGTGACTATTATAGACATGGTGATAGATAAAAAAAAAGATGAGA 916

Qy        918 GCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAA 977

Db        917 ATCAAGTTACTCAAATTAAGCTCCTTATGGATGTAAACATGGCTTG---TCTAAATGGAA 973

Qy        978 AAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACA 1037

Db        974 AAGAGAGAAATGAGGAAGAATGGAAGAAACTCTTCATAGAAGCTGGTTTCCAACACTATA 1033

Qy       1038 AAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGC 1097

Db       1034 AGATATCTCCTTTGACTGGATTTTTGTCTCTTATTGAGATCTATCCATAAACACTTTTGC 1093

Qy       1098 TGCAATATTCCATTTAGTAGTTAATTT 1124

Db       1094 TTTGATCATTCATCCATTCTATTGTTT 1120
```

```
RESULT 5
MSAF000976
LOCUS       MSAF000976              1227 bp    mRNA    linear   PLN 25-FEB-1998
DEFINITION  Medicago sativa 7-O-methyltransferase (7-IOMT(9)) mRNA, complete
            cds.
ACCESSION   AF000976
VERSION     AF000976.1  GI:2580583
KEYWORDS    .
SOURCE      Medicago sativa
  ORGANISM  Medicago sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 1227)
  AUTHORS   He,X.Z., Reddy,J.T. and Dixon,R.A.
  TITLE     Stress responses in alfalfa (Medicago sativa L). XXII. cDNA cloning
            and characterization of an elicitor-inducible isoflavone
            7-O-methyltransferase
  JOURNAL   Plant Mol. Biol. 36 (1), 43-54 (1998).
  MEDLINE   98145455
   PUBMED   9484461
REFERENCE   2  (bases 1 to 1227)
  AUTHORS   He,X.Z., Reddy,J.T., Guo,J.Z. and Dixon,R.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-1997) Plant Biology Divsion, The Samuel Roberts
            Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
FEATURES             Location/Qualifiers
     source          1. .1227
                     /organism="Medicago sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:3879"
     gene            1. .1227
                     /gene="7-IOMT(9)"
     CDS             24. .1082
                     /gene="7-IOMT(9)"
                     /codon_start=1
                     /product="7-O-methyltransferase"
                     /protein_id="AAC49927.1"
                     /db_xref="GI:2580584"
                     /translation="MASSINGRKPSEIFKAQALLYKHIYAPIDSMSLKWAVGMNIPNI
                     IHNHGKPISLSNLVSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEEESYALTVASEL
                     LVRGSDLCLAPMVECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEY
                     NTSFNDAMASDSKLINLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVF
                     DRPQVVENLSGSNNLTYVGGDMFTSIPNADAVLLKYILHNWTDKDCLRILKKCKEAVT
                     NDGKRGKVTIIDMVINEKKDENQVTQIKLLMDVNMACLNGKERNEEEWKKLFIEAGFQ
                     HYKISPLTGFLSLIEIYP"
BASE COUNT      405 a    182 c    237 g    403 t
ORIGIN
```

```
  Query Match             55.4%;  Score 642.2;  DB 8;  Length 1227;
  Best Local Similarity   75.8%;  Pred. No. 1.4e-135;
  Matches  839;  Conservative    0;  Mismatches  253;  Indels   15;  Gaps    3;

Qy         18 CTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACA 77

Db         28 CGTCATCAATTAATGGCAGAAAACCAAGTGAAATTTTCAAAGCACAAGCTTTATTATACA 87

Qy         78 AACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATAC 137

Db         88 AACACATATATGCCTTCATAGATTCCATGTCTCTTAAATGGGCTGTTGGAATGAACATAC 147

Qy        138 CCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAA 197

Db        148 CAAACATAATCCAC------AACCATGGCAAACCAATTTCTCTTTCAAACTTAGTTTCAA 201

Qy        198 TTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCAC 257

Db        202 TTCTTCAAGTTCCATCGTCGAAAATAGGTAACGTGCGGCGTCTCATGCGTTACCTTGCGC 261

Qy        258 ACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTG 317

Db        262 ACAACGGATTCTTCGAGATAATTACAAAAGAAGA------AGAGTCTTATGCTCTCACTG 315

Qy        318 CTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATT 377

Db        316 TTGCTTCAGAGCTTCTTGTTAGAGGCAGTGATCTTTGTTTAGCACCGATGGTTGAGTGTG 375

Qy        378 TTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAG 437

Db        376 TTCTTGATCCAACTCTTTCGGGTTCGTATCATGAGCTGAAGAAATGGATTTATGAGGAAG 435

Qy        438 ATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACC 497

Db        436 ATCTTACACTCTTTGGTGTTACTTTAGGATCTGGTTTTTGGGATTTTCTTGATAAAAATC 495

Qy        498 CTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGG 557

Db        496 CTGAATATAATACATCATTTAATGATGCAATGGCTAGTGATTCTAAATTGATAAACTTGG 555

Qy        558 CGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTG 617

Db        556 CATTGAGAGATTGTGATTTTGTGTTTGATGGATTGGAATCAATTGTGGATGTTGGTGGTG 615

Qy        618 GAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGT 677

Db        616 GAACTGGAACAACTGCTAAGATTATTTGTGAGACTTTTCCTAAGTTGAAATGTATTGTGT 675

Qy        678 TGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTG 737
```

```
Db        676 TTGATAGGCCACAAGTTGTAGAGAACTTATCTGGAAGCAATAATTTGACTTATGTTGGTG 735

Qy        738 GGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATA 797

Db        736 GGGACATGTTCACATCTATTCCTAATGCTGATGCAGTTTTGCTTAAGTATATTCTACATA 795

Qy        798 ATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTG 857

Db        796 ATTGGACTGATAAGGATTGCCTAAGGATACTGAAGAAATGTAAAGAAGCTGTTACAAATG 855

Qy        858 AAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGC 917

Db        856 ATGGGAAAAGAGGAAAAGTGACTATTATAGACATGGTGATAAATGAAAAAAAAGATGAGA 915

Qy        918 GCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAA 977

Db        916 ATCAAGTTACTCAAATTAAGCTCCTTATGGATGTAAACATGGCTTG---TCTAAATGGAA 972

Qy        978 AAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACA 1037

Db        973 AAGAGAGAAATGAGGAAGAATGGAAGAAACTCTTCATAGAAGCTGGTTTCCAACACTATA 1032

Qy       1038 AAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGC 1097

Db       1033 AGATATCTCCTTTGACTGGATTTTTGTCTCTTATTGAGATCTATCCATAAACACTTTTGC 1092

Qy       1098 TGCAATATTCCATTTAGTAGTTAATTT 1124

Db       1093 TTTGATCATTCATCCATTCTATTGTTT 1119
```

```
RESULT 6
AF023481
LOCUS       AF023481               1194 bp    mRNA    linear   PLN 07-DEC-1997
DEFINITION  Medicago sativa o-methytransferase (iomt2003) mRNA, complete cds.
ACCESSION   AF023481
VERSION     AF023481.1  GI:2565272
KEYWORDS    .
SOURCE      Medicago sativa
  ORGANISM  Medicago sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 1194)
  AUTHORS   He,X.Z. and Dixon,R.A.
  TITLE     A cDNA of an additional member of the
            isoflavone-O-methyltransferase (IOMT2) gene family in Medicago
            sativa (Accession No. AF023481) (PGR97-170)
  JOURNAL   Plant Physiol. 115, 1289 (1997)
REFERENCE   2  (bases 1 to 1194)
  AUTHORS   He,X.Z. and Dixon,R.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-1997) Plant Biology Divsion, The Samuel Roberts
            Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
FEATURES             Location/Qualifiers
     source          1. .1194
                     /organism="Medicago sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:3879"
     gene            1. .1194
                     /gene="iomt2003"
     CDS             28. .1059
                     /gene="iomt2003"
                     /function="methylates 7-hydroxyl-group of daidzein to
                     produce isoformononectin"
                     /codon_start=1
                     /product="o-methytransferase"
                     /protein_id="AAB88294.1"
                     /db_xref="GI:2565273"
                     /translation="MASSINGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNI
                     IQNHGKPISLSNLVSILQVPSSKIGNVRRLMRYLAHNGFFEIITKEEESYALTVASEL
                     LVRGSDLCLAPMVECVLDPTLSGSYHELKKWIYEEDLTLFGVTLGSGFWDFLDKNPEY
                     NTSFNDAMASDSKLINLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIVF
                     DRPQVVENLSGSNNLTYVGGDMFTSIPNADAVLLKDCLRILKKCKEAVTNDGKRGKVT
                     IIDMVIDEKKDENQVTQIKLLMDVNMACLNGKERNEEEWKKLFIEAGFQHYKISPLTG
                     FLSLIEIYP"
BASE COUNT      392 a    179 c    230 g    393 t
ORIGIN
```

```
  Query Match           51.2%;  Score 594.2;  DB 8;  Length 1194;
  Best Local Similarity 74.3%;  Pred. No. 1.1e-124;
  Matches  822;  Conservative    0;  Mismatches  243;  Indels   42;  Gaps    4;

Qy         18 CTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACA 77

Db         32 CTTCATCAATTAATGGCCGAAAACCAAGTGAAATTTTCAAAGCACAAGCTTTATTATACA 91

Qy         78 AACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATAC 137

Db         92 AACATATATATGCCTTCATAGATTCCATGTCTCTTAAATGGGCTGTTGAAATGAACATAC 151

Qy        138 CCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAA 197

Db        152 CAAACATAAT------CCAAAACCATGGCAAACCAATTTCTCTTTCAAACTTAGTTTCAA 205

Qy        198 TTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCAC 257

Db        206 TTCTTCAAGTTCCATCGTCGAAAATAGGTAACGTGCGGCGTCTCATGCGTTACCTCGCGC 265

Qy        258 ACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTG 317

Db        266 ACAACGGATTCTTCGAGATAATTACAAAAGAAGA------AGAGTCTTATGCTCTCACTG 319

Qy        318 CTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATT 377

Db        320 TTGCTTCAGAGCTTCTTGTTAGAGGCAGTGATCTTTGTTTAGCACCAATGGTTGAGTGTG 379

Qy        378 TTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAG 437

Db        380 TTCTTGATCCAACTCTTTCGGGTTCGTATCATGAGCTGAAGAAATGGATTTATGAGGAAG 439

Qy        438 ATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACC 497

Db        440 ATCTTACACTCTTTGGTGTTACTTTAGGATCTGGTTTTTGGGATTTTCTTGATAAAAATC 499

Qy        498 CTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGG 557

Db        500 CTGAATATAATACCTCATTTAATGATGCAATGGCTAGTGATTCTAAATTGATAAACTTGG 559

Qy        558 CGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTG 617

Db        560 CATTGAGAGATTGTGATTTTGTGTTTGATGGATTGGAATCAATTGTGGATGTTGGTGGTG 619

Qy        618 GAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGT 677

Db        620 GAACTGGAACAACTGCTAAGATTATTTGTGAGACTTTTCCTAAGTTGAAATGTATTGTGT 679

Qy        678 TGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTG 737

Db        680 TTGATAGGCCACAAGTTGTAGAGAACTTATCTGGAAGCAATAATTTGACTTATGTTGGTG 739

Qy        738 GGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATA 797

Db        740 GGGACATGTTCACATCTATTCCTAATGCTGATGCAGTTTTGCTTAAG------------- 786

Qy        798 ATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTG 857

Db        787 -------------GATTGCCTAAGGATACTGAAGAAATGTAAAGAAGCTGTTACAAATG 832

Qy        858 AAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGC 917

Db        833 ATGGGAAAAGAGGAAAAGTGACTATTATAGACATGGTGATAGATGAAAAAAAAGATGAGA 892

Qy        918 GCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAA 977
```

```
Db        893 ATCAAGTTACTCAAATTAAGCTCCTTATGGATGTAAACATGGCTTG---TCTAAATGGAA 949

Qy        978 AAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACA 1037

Db        950 AAGAGAGAAATGAGGAAGAATGGAAGAAACTCTTCATAGAAGCTGGTTTCCAACACTATA 1009

Qy       1038 AAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGC 1097

Db       1010 AGATATCTCCTTTGACTGGATTTTTGTCTCTTATTGAGATCTATCCATAAACACTTTTGC 1069

Qy       1098 TGCAATATTCCATTTAGTAGTTAATTT 1124

Db       1070 TTTGATCATTCATCCATTCTATTGTTT 1096
```

```
RESULT 7
AB091684
LOCUS       AB091684                1349 bp    mRNA    linear   PLN 04-MAR-2003
DEFINITION  Glycyrrhiza echinata HI4'OMT mRNA for S-adenosyl-L-methionine:
            2,7,4'-trihydroxyisoflavanone 4'-O-methyltransferase, complete cds.
ACCESSION   AB091684
VERSION     AB091684.1  GI:28804591
KEYWORDS    .
SOURCE      Glycyrrhiza echinata
  ORGANISM  Glycyrrhiza echinata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae;
            Glycyrrhiza.
REFERENCE   1
  AUTHORS   Akashi,T., Sawada,Y., Shimada,N., Sakurai,N., Aoki,T. and Ayabe,S.
  TITLE     cDNA cloning and biochemical characterization of
            s-adenosyl-L-methionine: 2,7,4'-trihydroxyisoflavanone
            4'-o-methyltransferase, a critical enzyme of the legume
            isoflavonoid phytoalexin pathway
  JOURNAL   Plant Cell Physiol. 44 (2), 103-112 (2003)
  MEDLINE   22497945
   PUBMED   12610212
REFERENCE   2  (bases 1 to 1349)
  AUTHORS   Akashi,T., Aoki,T. and Ayabe,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-SEP-2002) Shin-ichi Ayabe, Nihon University,
            Department of Applied Biological Sciences; Kameino 1866, Fujisawa,
            Kanagawa 252-8510, Japan (E-mail:ayabe@brs.nihon-u.ac.jp,
            Tel:81-466-84-3703, Fax:81-466-80-1141)
FEATURES             Location/Qualifiers
     source          1. .1349
                     /organism="Glycyrrhiza echinata"
                     /mol_type="mRNA"
                     /db_xref="taxon:46348"
                     /cell_line="Ak-1"
                     /note="cloning vector: lambda ZapII-cultured cells"
     gene            1. .1349
                     /gene="HI4'OMT"
     CDS             35. .1138
                     /gene="HI4'OMT"
                     /codon_start=1
                     /product="S-adenosyl-L-methionine:
                     2,7,4'-trihydroxyisoflavanone 4'-O-methyltransferase"
                     /protein_id="BAC58011.1"
                     /db_xref="GI:28804592"
                     /translation="MAFSTNGSEEIELYHAQIHLYKHVYNFVSSMALKSAMELGIADV
                     IHNHGKPITLPELASALKLHPSKVGILYRFLRLLTHNGFFAKTTVPSQNGKDGEEEEE
                     TAYALTPPSKLLVKGKPTCLASIVRGALHPSSLDMWRSSEKWFKEDKELTLFESATGE
                     SFWDFLNKDSESGTLSMFQEAMAADSQMFKLALKECRHVFEGLESLVDVGGGTGGVTK
                     LIHEEFPHLKCTVFDQPQVVGNLSGNENLKFVGGDMFKSIPPADAVLLKWVLHDWNDE
                     LSLKILKNSKEAISGKGKEGKVIIIDISIDEASGDRELTELQLDYDLVMLTMFNGKER
                     EKKEWEKLISDAGFSSYKITPICGFKSLIEVFP"
BASE COUNT      410 a    280 c    293 g    366 t
ORIGIN
```

Query Match 29.1%; Score 337.6; DB 8; Length 1349;

Best Local Similarity 59.9%; Pred. No. 2.5e-66;
Matches 659; Conservative 0; Mismatches 399; Indels 42; Gaps 4;

```
Qy         26 AACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTG 85

Db         47 ACCAATGGCAGTGAAGAAATTGAACTGTACCATGCTCAAATCCATCTCTACAAGCATGTC 106

Qy         86 CTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATA 145

Db        107 TACAACTTTGTAAGTTCCATGGCTCTCAAATCCGCTATGGAGCTCGGCATAGCAGATGTA 166

Qy        146 ATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAA 205

Db        167 ATCCACAA------CCATGGCAAACCCATCACTCTTCCCGAATTAGCCTCAGCCTTGAAA 220

Qy        206 GTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGA 265

Db        221 CTCCACCCTTCCAAAGTCGGTATCCTCTACCGGTTCCTGCGTCTTCTAACACACAACGGG 280

Qy        266 TTCTTTGAGATAGTAAGAATCC--------------------------ATGACAACATA 298

Db        281 TTCTTCGCCAAAACAACAGTTCCATCACAGAACGGCAAAGATGGGGAAGAAGAAGAAGAA 340

Qy        299 GAAGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTA 358

Db        341 ACAGCATATGCTTTAACACCCCCATCAAAGCTTCTCGTTAAAGGAAAGCCAACTTGTTTG 400

Qy        359 GCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAG 418

Db        401 GCATCAATTGTTCGTGGAGCACTTCATCCAAGTTCTCTGGACATGTGGCGATCCTCTGAG 460

Qy        419 AGGTGGGTTCATGAGG---AAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTC 475

Db        461 AAATGGTTCAAAGAGGACAAGGAGCTCACCCTGTTTGAGAGCGCAACAGGGGAGAGTTTC 520

Qy        476 TGGGACTTTATCAATAAAGACCCTGCATATAA------CAAGTCATTCAATGAGGCAATG 529

Db        521 TGGGACTTTCTCAACAAGGATTCTGAATCTGGTACACTCAGCATGTTTCAAGAAGCCATG 580

Qy        530 GCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGA 589

Db        581 GCTGCTGATTCTCAAATGTTCAAGCTTGCCCTCAAGGAGTGTAGGCATGTGTTTGAGGGT 640

Qy        590 CTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAG 649

Db        641 TTGGAGTCTCTTGTTGATGTTGGAGGTGGGACTGGTGGTGTCACAAAACTCATCCATGAA 700

Qy        650 GCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCA 709

Db        701 GAATTTCCTCACCTCAAATGCACAGTTTTTGACCAACCACAGGTTGTGGGTAACTTGTCT 760

Qy        710 GGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGAT 769

Db        761 GGGAATGAAAATTTGAAATTTGTTGGTGGAGATATGTTCAAGTCTATCCCTCCTGCAGAT 820

Qy        770 GCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTA 829

Db        821 GCAGTTTTACTCAAGTGGGTTCTGCATGATTGGAACGATGAACTCTCCTTGAAGATATTG 880

Qy        830 GAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGAT 889

Db        881 AAGAACAGCAAAGAAGCTATTTCTGGGAAAGGGAAAGAAGGGAAGGTGATTATCATAGAC 940

Qy        890 ACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGAT 949

Db        941 ATATCGATTGATGAAGCAAGTGGTGATCGCGAATTGACTGAGTTGCAACTGGACTATGAC 1000

Qy        950 GTACACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTC 1009

Db       1001 TTGGTGATGCTGACTATGTTCAATGGAAAAGAAAGAGAGAAGAAAGAGTGGGAGAAACTC 1060

Qy       1010 TTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTT 1069
```

```
Db       1061 ATATCCGACGCAGGGTTCAGCAGCTACAAGATTACCCCCATTTGTGGCTTCAAGTCCCTC 1120

Qy       1070 ATTGAGATCTATCCTTGAAT 1089

Db       1121 ATTGAAGTTTTTCCTTAATT 1140


RESULT 8
PSU69554
LOCUS       PSU69554                1337 bp    mRNA    linear   PLN 29-OCT-1997
DEFINITION  Pisum sativum 6a-hydroxymaackiain methyltransferase (hmm6) mRNA,
            complete cds.
ACCESSION   U69554
VERSION     U69554.1  GI:1568636
KEYWORDS    .
SOURCE      Pisum sativum (pea)
  ORGANISM  Pisum sativum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae;
            Pisum.
REFERENCE   1  (bases 1 to 1337)
  AUTHORS   Wu,Q., Preisig,C.L. and VanEtten,H.D.
  TITLE     Isolation of the cDNAs encoding (+)6a-hydroxymaackiain
            3-O-methyltransferase, the terminal step for the synthesis of the
            phytoalexin pisatin in Pisum sativum
  JOURNAL   Plant Mol. Biol. 35 (5), 551-560 (1997)
  MEDLINE   98009990
   PUBMED   9349277
REFERENCE   2  (bases 1 to 1337)
  AUTHORS   Wu,Q., Preisig,C.L. and VanEtten,H.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-SEP-1996) Plant Pathology, University of Arizona,
            Forbes Building, Room 204, Tucson, AZ 85721, USA
FEATURES             Location/Qualifiers
     source          1..1337
                     /organism="Pisum sativum"
                     /mol_type="mRNA"
                     /strain="Alaska"
                     /db_xref="taxon:3888"
     gene            1..1337
                     /gene="hmm6"
     CDS             14..1096
                     /gene="hmm6"
                     /function="pisatin biosynthesis;  catalyzes the
                     methylation of (+)6a-hydroxymaackiain to (+)pisatin"
                     /note="methyltransferase"
                     /codon_start=1
                     /product="6a-hydroxymaackiain methyltransferase"
                     /protein_id="AAC49856.1"
                     /db_xref="GI:1568637"
                     /translation="MDFSTNGSEESELYHAQIHLYKHVYNFVSSMALKSAMELGIADA
                     IHNHGKPMTLPELSSSLKLHPSKVNILYRFLRLLTHNGFFAKTTVKSNEGEEETAYVL
                     TPSSKLLVSGKSTCLSSLVKGALHPSSLDMWGVSKKWFHEDKEQTLFECATGENYWDF
                     LNKDSDSLSMFQDAMAADSRLFKLAIQENKHVFEGLESLVDVAGGTGGVAKLIHEAFP
                     HIKCTVFDQPQVVGNLTGNENLNFVGGDMFKSVPSADAVLLKWVLHDWNDELSLKILK
                     NSKEAISHKGKDGKVIIIDISIDENSDDRGLTELQLEYDVVMLTMFLGKERTKKEWEK
                     LIYDAGFSRYKITPICGFKSLIEVYP"
BASE COUNT      450 a    204 c    266 g    417 t
ORIGIN

  Query Match             27.7%;  Score 321.6;  DB 8;  Length 1337;
  Best Local Similarity   58.5%;  Pred. No. 1.1e-62;
  Matches  627;  Conservative    0;  Mismatches  424;  Indels   21;  Gaps    3;

Qy         29 AATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTT 88

Db         29 AACGGTTCTGAAGAAAGTGAACTCTATCATGCTCAAATCCATTTATACAAACATGTATAC 88

Qy         89 GGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATC 148

Db         89 AACTTTGTAAGTTCCATGGCACTCAAATCCGCCATGGAACTAGGCATAGCTGATGCAATT 148

Qy        149 CACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTC 208

Db        149 ------CATAACCATGGAAAACCAATGACTCTCCCTGAATTATCTTCATCTTTGAAACTT 202

Qy        209 CCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTC 268

Db        203 CACCCCTCTAAAGTTAACATCCTTTATCGATTCTTACGTCTGTTAACACACAATGGTTTC 262

Qy        269 TTTGAGATAGTAAGAATCCATGACAACATAG-----------AAGCATATGCTCTCACT 316

Db        263 TTTGCTAAAACAACAGTGAAAAGCAATGAAGGAGAAGAAGAAACAGCCTATGTTCTCACA 322

Qy        317 GCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTAT 376

Db        323 CCTTCTTCAAAGCTTCTAGTAAGTGGTAAATCAACATGTTTATCATCACTTGTTAAAGGA 382

Qy        377 TTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGG-- 434

Db        383 GCACTTCATCCAAGTTCTTTAGACATGTGGGGTGTTTCAAAGAAATGGTTCCATGAAGAT 442

Qy        435 -AAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAA 493

Db        443 AAGGAACAAACACTTTTTGAGTGTGCAACTGGGGAGAACTATTGGGATTTTCTTAACAAA 502

Qy        494 GACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAAC 553

Db        503 GATTCTGACAGTTTGAGTATGTTTCAAGATGCTATGGCGGCTGATTCGAGATTGTTTAAG 562

Qy        554 TTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGT 613

Db        563 CTTGCTATTCAAGAGAATAAGCACGTGTTTGAGGGTTTGGAGTCTTTGGTTGATGTGGCA 622

Qy        614 GGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATG 673

Db        623 GGTGGAACTGGTGGTGTTGCAAAACTCATTCATGAAGCATTTCCTCACATCAAATGCACT 682

Qy        674 GTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTT 733

Db        683 GTTTTTGATCAGCCACAGGTTGTTGGTAACTTAACTGGAAATGAAAATTTGAACTTTGTT 742

Qy        734 GGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTA 793

Db        743 GGTGGAGATATGTTCAAGTCTGTCCCTTCTGCTGATGCTGTTTTACTCAAGTGGGTTCTG 802

Qy        794 CATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCA 853

Db        803 CATGATTGGAATGATGAACTATCTTTAAAGATATTGAAGAACAGTAAAGAAGCAATTTCA 862

Qy        854 GGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGAT 913

Db        863 CACAAAGGGAAAGATGGAAAAGTGATTATCATAGACATATCAATTGATGAAAACAGTGAT 922

Qy        914 GAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAAT 973

Db        923 GATCGTGGATTAACTGAGTTGCAATTGGAGTATGATGTAGTGATGCTTACAATGTTTCTT 982

Qy        974 GGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGC 1033

Db        983 GGTAAAGAAAGAACAAAGAAAGAATGGGAGAAACTCATATACGATGCAGGTTTCAGCCGC 1042

Qy       1034 TACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTT 1085

Db       1043 TACAAAATTACTCCTATATGTGGCTTCAAGTCCCTCATTGAAGTTTATCCTT 1094


RESULT 9
AC136140/c
LOCUS       AC136140              111587 bp    DNA     linear   HTG 07-JUN-2003
DEFINITION  Medicago truncatula clone mth2-11a20, WORKING DRAFT SEQUENCE, 3
            ordered pieces.
ACCESSION   AC136140
VERSION     AC136140.8  GI:31455653
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
```

```
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 111587)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Medicago truncatula BAC Clone mth2-11a20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 111587)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-2002) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 111587)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Jun 7, 2003 this sequence version replaced gi:30061413.
            -------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1     9681: contig of 9681 bp in length
            *     9682     9781: gap of unknown length
            *     9782    46738: contig of 36957 bp in length
            *    46739    46838: gap of unknown length
            *    46839   111587: contig of 64749 bp in length.
FEATURES             Location/Qualifiers
     source          1. .111587
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /clone="mth2-11a20"
                     /clone_lib="Medicago truncatula BAC library H2"
                     /note="This is one of two clones in the same well from
                     mth2-11a20"
BASE COUNT    37516 a  18138 c  18553 g  37160 t    220 others
ORIGIN

  Query Match           27.5%;  Score 318.8;  DB 2;  Length 111587;
  Best Local Similarity   58.8%;  Pred. No. 2.7e-62;
  Matches  699;  Conservative    0;  Mismatches  347;  Indels  142;  Gaps    3;

Qy         27 ACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGC 86
Db     104606 AGAATGAGCACAAAGCAAGTGAACTCTTTGAAGTTCAAGCTCACTTGTACAATCAGTTAT 104547

Qy         87 TTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAA 146
Db     104546 TTAGTTTCTTGAGACCCGTGTCGATCAAGTGGGCTGTTGAACTTGGCATACCAGATATAA 104487

Qy        147 TCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAG 206
Db     104486 T------TCAAAATCATGGCAAACCTATTACTCTTCCCGAGCTTGTCTCGGCTCTTCGAA 104433

Qy        207 TCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGAT 266
Db     104432 TTCCAGAAGCAAAAGCCGGTTGTGTTCACAGTGTCATGCGTTTATTGGCGCACAATAAAA 104373

Qy        267 TCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTGCTGCTTCAG 326
Db     104372 TCTTCGCAATTGTGAAGATAGACGACAACAAAGAAGCATATGCTCTTTCTCCAACATCAA 104313

Qy        327 AGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAAC 386
Db     104312 AGCTTCTTGTGAAAGGAACTGATCATTGTTTATCTTCAATGGTTAAGTTGCTAACAAACC 104253

Qy        387 CAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAG 446
Db     104252 CAACTCGCGTCGAAAAACATTACCAATTGAGCACATGGACATCAATGGAGGACCTCACAA 104193

Qy        447 TATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATA 506
Db     104192 TCTTTGAG------------ACTCCATTATGGGATTTAATTTTACAAAATCCGACTCATT 104145

Qy        507 ACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAG 566
Db     104144 CCAAGTTATTCAATGATGCTATGGAAAGTGATTCTAATATGGTGAGGTTTGCAATGAGTG 104085

Qy        567 ATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAA 626
Db     104084 ATTGCAAATCAGTTTTTGAGGGCTTAACTTCGTTGGTTGATGTTGGAGGTGGAACCGGAA 104025

Qy        627 TCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTC 686
Db     104024 ACACTGTTAAAATTATTTGTGAGGCGTTTCCTACGTTAAAATGTATAGTGTTCGACCTTC 103965

Qy        687 CAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGT 746
Db     103964 CAAACGTGGTAGAAGGATTAACGGGAAATAACTATCTGAGTTTTGTAGGCGGAAACATGT 103905

Qy        747 TTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTA---------------------- 782
Db     103904 TTGAATCTATCCCTCAAGCTGATGCAATTTTACTCAAGGTATGTATGCAGTTTTGTAAGT 103845

Qy        783 ---------------------------------------------------------- 782
Db     103844 CATGCATTATTAAATAGCTTAGTTTCCTGATTCTTGATACATTGAATCCTAAAAATACCA 103785

Qy        783 ----------------------------------------AGTTGGTTTTACATAATTGG 802
Db     103784 TCAAGTCTTTCTCATTTTTATTTTTGTTTTCCACCTTTGTAGTGGATTTTGCATAACTGG 103725

Qy        803 AATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGC 862
Db     103724 AATGATGATGACTGTGTCAAGATACTTAGAAATTGTAAAGAAGCTGTATCAAGGAAAAAC 103665

Qy        863 AAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAA 922
Db     103664 AAAGAAGGAAAAGTGATCATAATAGATATTGTGATAAATGAAGAGCAAGATGAGCATGAA 103605

Qy        923 GTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAG 982
Db     103604 ATGACAGAACTCAAACTCTTCTTTGACATAAGTATGATGGCTAATTTCGGTTCAACAGAA 103545

Qy        983 AGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATA 1042
Db     103544 AGAGATGAAAAAGCATGGGAGAAAATCATCACTGAAGCAGGATTTACATCTTACAAAATA 103485

Qy       1043 TCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATA 1090
Db     103484 TTTCATATATTTGGTTTCAAATCACTTATTGAACTTTGTGTTTAATTA 103437

RESULT 10
AB091686
LOCUS       AB091686               1363 bp    mRNA    linear   PLN 04-MAR-2003
DEFINITION  Lotus japonicus HI4'OMT mRNA for S-adenosyl-L-methionine:
```

```
            2,7,4'-trihydroxyisoflavanone 4'-O-methyltransferase, complete cds.
ACCESSION   AB091686
VERSION     AB091686.1  GI:28804595
KEYWORDS    .
SOURCE      Lotus japonicus
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
   PUBMED   10819328
REFERENCE   2
  AUTHORS   Akashi,T., Sawada,Y., Shimada,N., Sakurai,N., Aoki,T. and Ayabe,S.
  TITLE     cDNA cloning and biochemical characterization of
            s-adenosyl-L-methionine: 2,7,4'-trihydroxyisoflavanone
            4'-o-methyltransferase, a critical enzyme of the legume
            isoflavonoid phytoalexin pathway
  JOURNAL   Plant Cell Physiol. 44 (2), 103-112 (2003)
  MEDLINE   22497945
   PUBMED   12610212
REFERENCE   3  (bases 1 to 1363)
  AUTHORS   Akashi,T., Aoki,T. and Ayabe,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-SEP-2002) Shin-ichi Ayabe, Nihon University,
            Department of Applied Biological Sciences; Kameino 1866, Fujisawa,
            Kanagawa 252-8510, Japan (E-mail:ayabe@brs.nihon-u.ac.jp,
            Tel:81-466-84-3703, Fax:81-466-80-1141)
FEATURES             Location/Qualifiers
     source          1. .1363
                     /organism="Lotus japonicus"
                     /mol_type="mRNA"
                     /strain="MG-20 Miyakojima"
                     /db_xref="taxon:34305"
                     /note="cloning vector: lambda ZapII"
     gene            1. .1363
                     /gene="HI4'OMT"
     misc_feature    1. .296
                     /gene="HI4'OMT"
                     /note="Almost identical to entire nucleotide sequence of
                     AV407445."
     CDS             54. .1151
                     /gene="HI4'OMT"
                     /codon_start=1
                     /product="S-adenosyl-L-methionine:
                     2,7,4'-trihydroxyisoflavanone 4'-O-methyltransferase"
                     /protein_id="BAC58013.1"
                     /db_xref="GI:28804596"
                     /translation="MDFSSSNGSEDTELSQAQIHLYKHVYNFVSSMALKSAMELGIAD
                     VIHSHGKPITLPELATALNLRPSKIGVLHRFLRLLTHNGFFAKTTVSRGEGAEEETAY
                     GLTPPSKLLVKSNSTCLAPIVKGALHPSSLDMWRSSKKWFLEDNEELTLFESATGESF
                     WEFLNKETESDTLSMFQEAMAADSHMFKLALKECKHVFEGLGSLVDVAGGRGGVTKLI
                     REAFPHVKCTVFDQPQVVANLTGDENLNFVGGDMFKSVPPADAVLLKWVLHDWNDELS
                     LKILKNCKEAISGRGKEGKVIIIDISIDETSDDRELTELKLDYDLVMLTMFNGKEREK
                     KEWEKLIYDAGFSSYKITPICGFKSLIEVFP"
BASE COUNT      385 a    288 c    322 g    368 t
ORIGIN
```

```
  Query Match             27.4%;  Score 318;  DB 8;  Length 1363;
  Best Local Similarity   58.4%;  Pred. No. 7.2e-62;
  Matches  649;  Conservative    0;  Mismatches  430;  Indels   33;  Gaps    4;

Qy         11 ATGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTC 70
Db         54 ATGGATTTCAGCTCCAGCAATGGCAGTGAAGACACTGAACTTTCCCAAGCTCAGATTCAC 113

Qy         71 TTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTT 130

Db        114 CTCTACAAACACGTCTACAACTTCGTCAGCTCCATGGCTCTCAAATCCGCCATGGAACTA 173

Qy        131 GACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTG 190

Db        174 GGCATCGCCGACGTGATCCAC------AGCCATGGAAAACCCATCACCCTCCCTGAGTTA 227

Qy        191 GTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTAT 250

Db        228 GCCACTGCCTTGAACCTCCGCCCTTCCAAAATCGGCGTCCTCCACCGCTTCCTCCGCCTG 287

Qy        251 CTAGCACACAATGGATTCTTTGAGA--------------TAGTAAGAATCCATGACAAC 295

Db        288 CTAACCCACAACGGTTTCTTCGCCAAAACCACCGTGTCCCGAGGTGAAGGAGCAGAGGAA 347

Qy        296 ATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGT 355

Db        348 GAAACAGCGTACGGTCTCACACCGCCATCGAAGCTTCTGGTGAAAAGCAACTCAACGTGT 407

Qy        356 TTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTG 415

Db        408 TTAGCTCCAATTGTGAAGGGAGCACTTCATCCAAGCTCGTTGGACATGTGGCGATCGTCG 467

Qy        416 AAGAGGTGGGTTCATGAG------GAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACA 469

Db        468 AAGAAATGGTTCTTGGAGGACAATGAAGAACTGACTCTGTTTGAGAGCGCAACAGGGGAG 527

Qy        470 CCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCA------ATGAG 523

Db        528 AGCTTCTGGGAGTTTCTGAACAAGGAAACGGAGTCTGATACGCTGAGCATGTTTCAGGAG 587

Qy        524 GCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTT 583

Db        588 GCTATGGCGGCGGATTCTCACATGTTTAAGCTTGCTTTGAAGGAGTGTAAGCATGTGTTT 647

Qy        584 GAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATC 643

Db        648 GAGGGTTTGGGTTCTCTTGTTGATGTTGCAGGGGGAAGAGGGGGTGTTACCAAACTGATT 707

Qy        644 TGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAAT 703

Db        708 CGTGAAGCGTTCCCTCATGTGAAGTGTACGGTTTTTGATCAGCCTCAGGTTGTGGCTAAC 767

Qy        704 TTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAG 763

Db        768 TTGACTGGAGATGAGAATTTGAACTTTGTTGGTGGTGATATGTTCAAGTCTGTCCCTCCT 827

Qy        764 GCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAG 823

Db        828 GCAGATGCTGTTCTGTTGAAGTGGGTTCTGCATGACTGGAATGATGAGCTCTCCTTAAAG 887

Qy        824 ATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTC 883

Db        888 ATATTGAAGAACTGCAAAGAAGCTATTTCAGGGAGAGGGAAAGAAGGCAAAGTGATTATC 947

Qy        884 ATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTT 943

Db        948 ATAGACATATCAATTGATGAAACAAGTGATGACAGAGAACTGACTGAGTTGAAATTGGAC 1007

Qy        944 ATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAG 1003

Db       1008 TATGACTTGGTCATGCTGACTATGTTCAATGGAAAAGAAAGAGAGAAGAAAGAATGGGAG 1067

Qy       1004 AAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTG 1063

Db       1068 AAGCTCATATATGATGCAGGCTTCAGCAGCTACAAGATTACACCCATATGTGGCTTCAAG 1127

Qy       1064 TCTCTTATTGAGATCTATCCTTGAATACTGAC 1095

Db       1128 TCTCTCATTGAAGTTTTTCCTTAATTAATTAC 1159


RESULT 11
PAU82011
```

```
LOCUS       PAU82011                1257 bp    mRNA    linear   PLN 25-SEP-1997
DEFINITION  Prunus armeniaca methyltransferase mRNA, complete cds.
ACCESSION   U82011
VERSION     U82011.1  GI:2282585
KEYWORDS    .
SOURCE      Prunus armeniaca (apricot)
  ORGANISM  Prunus armeniaca
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
REFERENCE   1  (bases 1 to 1257)
  AUTHORS   Mbeguie-A-Mbeguie,D., Gomez,R.-M. and Fils-Lycaon,B.
  TITLE     Sequence of an O-Methyltransferase from Apricot Fruit (Accession
            No. U82011). Gene Expression During Fruit Ripening (PGR97-118)
  JOURNAL   Plant Physiol. 114, 1569 (1997)
REFERENCE   2  (bases 1 to 1257)
  AUTHORS   Mbeguie A Mbeguie,D., Gomez,R.-M. and Fils-Lycaon,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-DEC-1996) Technologie des Produits Vegetaux, INRA,
            Site AGROPARC, Domaine Saint Paul, Avignon 84914 cedex, France
FEATURES             Location/Qualifiers
     source          1. .1257
                     /organism="Prunus armeniaca"
                     /mol_type="mRNA"
                     /strain="Bergeron"
                     /db_xref="taxon:36596"
                     /tissue_type="fruit mesocarp and exocarp"
                     /dev_stage="ripe"
     CDS             10. .1074
                     /codon_start=1
                     /product="methyltransferase"
                     /protein_id="AAB71213.1"
                     /db_xref="GI:2282586"
                     /translation="MGSVRASHELLQAQAHIWNHIFSFINSLSLKCAVQLDIPDVIQK
                     HGQPMTLSELVSALPISPTKAHFIPRLMRILVHSGFFAKESLSGCGEQGYILTDASAL
                     LLKDNPMSARPFLLAMLSPILTDPYQYLTTWFQNDNPTPFHVVNGMTCWDYVNQDPTL
                     AHFFNDAMASDAQLISSLVIDDCKEVFQGVDSLVDVGGGTGTVAKSIADAFPHMKCTV
                     LDLPHVVADLKGSKNLEYVAGDMFEAVPAADAIFLKWILHDWSDEECVKILERCKAAV
                     TREGKKGKVIIVEMTVENKNTDKESGETQLFFDMHMMVMSTGKERNEKEWAKLFSDAG
                     FSQYKITPLFGFKSLIEVYP"
BASE COUNT      362 a    252 c    285 g    358 t
ORIGIN
```

```
  Query Match           26.3%;  Score 305.6;  DB 8;  Length 1257;
  Best Local Similarity   57.5%;  Pred. No. 4.8e-59;
  Matches  611;  Conservative    0;  Mismatches  439;  Indels   12;  Gaps    3;

Qy         46 TGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAA 105
Db         33 TGAACTACTCCAAGCACAAGCCCACATTTGGAACCACATTTTCAGCTTCATAAATTCTTT 92

Qy        106 GTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGG 165
Db         93 GTCCCTCAAATGTGCAGTTCAATTAGATATCCCAGATGTCATCCAAAAA------CATGG 146

Qy        166 CCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCG 225
Db        147 CCAACCCATGACTCTTTCCGAGCTTGTCTCCGCCCTACCAATTTCCCCAACAAAAGCTCA 206

Qy        226 TCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAAT 285
Db        207 TTTCATCCCACGCCTCATGCGAATCTTAGTCCACTCCGGCTTCTTTGCCAAAGAAAGTCT 266

Qy        286 CCATG---ACAACATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAG 342
Db        267 AAGTGGCTGCGGCGAACAAGGTTATATATTAACAGACGCCTCAGCACTCCTCCTGAAAGA 326

Qy        343 CAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGC 402
Db        327 CAATCCCATGAGCGCAAGGCCCTTCTTACTTGCCATGCTCAGTCCCATCTTAACCGATCC 386

Qy        403 ATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTT 462
Db        387 ATATCAGTATTTGACCACCTGGTTCCAAAATGACAATCCTACGCCGTTTCACGTGGTTAA 446

Qy        463 AGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGA 522
Db        447 TGGGATGACATGTTGGGATTATGTGAACCAGGACCCGACTCTTGCACATTTTTTCAATGA 506

Qy        523 GGCAATGGCTTGTGATTCTCA---GATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGT 579
Db        507 CGCCATGGCTAGCGATGCCCAGTTGATCTCGAGCTTGGTGATTGATGACTGCAAGGAGGT 566

Qy        580 CTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGAT 639
Db        567 GTTTCAAGGAGTGGATTCGTTGGTTGACGTGGGCGGTGGGACGGGAACTGTGGCCAAGTC 626

Qy        640 TATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGA 699
Db        627 CATTGCTGATGCCTTCCCTCATATGAAATGCACTGTACTTGATCTCCCTCATGTAGTTGC 686

Qy        700 AAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCC 759
Db        687 TGACCTAAAAGGGAGTAAGAACTTGGAATATGTTGCAGGGGACATGTTTGAGGCTGTTCC 746

Qy        760 CAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCAT 819
Db        747 TGCGGCCGATGCTATTTTTTTGAAGTGGATATTGCACGACTGGAGTGATGAAGAATGCGT 806

Qy        820 GAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGT 879
Db        807 GAAAATACTCGAGCGATGTAAAGCGGCAGTTACAAGAGAGGGCAAGAAAGGCAAGGTGAT 866

Qy        880 TGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCT 939
Db        867 TATCGTAGAAATGACGGTGGAGAACAAGAATACAGATAAGGAATCTGGGGAAACACAACT 926

Qy        940 CCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTG 999
Db        927 TTTCTTCGATATGCATATGATGGTTATGAGTACAGGAAAAGAAAGGAATGAGAAAGAATG 986

Qy       1000 GAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATA 1059
Db        987 GGCTAAGCTCTTTTCTGACGCAGGTTTCAGCCAGTATAAGATTACTCCCCTGTTTGGGTT 1046

Qy       1060 TTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGCTGCA 1101
Db       1047 TAAATCTCTCATTGAGGTTTATCCTTGATTTATAATTAAGCA 1088
```

```
RESULT 12
AF502434
LOCUS       AF502434                1253 bp    mRNA    linear   PLN 14-AUG-2002
DEFINITION  Rosa hybrid cultivar orcinol O-methyltransferase (OOMT2) mRNA,
            complete cds.
ACCESSION   AF502434
VERSION     AF502434.1  GI:20514368
KEYWORDS    .
SOURCE      Rosa hybrid cultivar
  ORGANISM  Rosa hybrid cultivar
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
REFERENCE   1  (bases 1 to 1253)
  AUTHORS   Lavid,N., Wang,J., Shalit,M., Guterman,I., Bar,E., Beuerle,T.,
            Menda,N., Shafir,S., Zamir,D., Adam,Z., Vainstein,A., Weiss,D.,
            Pichersky,E. and Lewinsohn,E.
  TITLE     O-methyltransferases involved in the biosynthesis of volatile
            phenolic derivatives in rose petals
  JOURNAL   Plant Physiol. 129 (4), 1899-1907 (2002)
  MEDLINE   22167322
   PUBMED   12177504
REFERENCE   2  (bases 1 to 1253)
  AUTHORS   Lavid,N., Wang,J., Shalit,M., Guterman,I., Bar,E., Beuerle,T.,
            Menda,N., Shafir,S., Zamir,D., Adam,Z., Vainstein,A., Weiss,D.,
            Pichersky,E. and Lewinsohn,E.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-2002) Vegetables Crops, Newe Ya'ar Research
            Center, Agricultural Research Organization, P.O. Box 1021, Ramat
            Yishay 30095, Israel
FEATURES             Location/Qualifiers
     source          1. .1253
                     /organism="Rosa hybrid cultivar"
                     /mol_type="mRNA"
                     /cultivar="Fragrant Cloud"
                     /isolate="fc0895"
                     /db_xref="taxon:128735"
                     /tissue_type="petals"
     gene            1. .1253
                     /gene="OOMT2"
     CDS             8. .1108
                     /gene="OOMT2"
                     /function="catalyzes the S-adenosylmethionine-dependent
                     O-methylation of 3-methoxy, 5 hydroxytoluene to form
                     3,5-dimethoxytoluene (orcinol dimethyl ether)"
                     /codon_start=1
                     /product="orcinol O-methyltransferase"
                     /protein_id="AAM23005.1"
                     /db_xref="GI:20514369"
                     /translation="MERLNSFKHLNQKWSNGEHSNELLHAQAHIWNHIFSFINSMSLK
                     SAIQLGIPDIINKHGPMTLSELTSALPIHPTKSHSVYRLMRILVHSGFFAKKKLSKTD
                     EEGYTLTDASQLLLKDHPLSLTPFLTAMLDPVLTTPWNYLSTWFQNEDPTPFDTAHGM
                     TFWDYGNHQPSIAHLFNDAMASDARLVTSVIIDDCKGVFEGLESLVDVGGGTGTVAKA
                     IADAFPHIECTVLDLPHVVADLQGSKNLKYTGGDMFEAVPPADTVLLKWILHDWNDEE
                     CIKILKRSRVAITSKDKKGKVIIIDMMMENQKGDEESIETQLFFDMLMMALVRGQERN
                     EKEWAKLFTDAGFSDYKITPILGLRSLIEVYP"
BASE COUNT      389 a    270 c    265 g    329 t
ORIGIN
```

Query Match 26.1%; Score 302.8; DB 8; Length 1253;
Best Local Similarity 57.3%; Pred. No. 2.1e-58;
Matches 613; Conservative 0; Mismatches 442; Indels 15; Gaps 3;

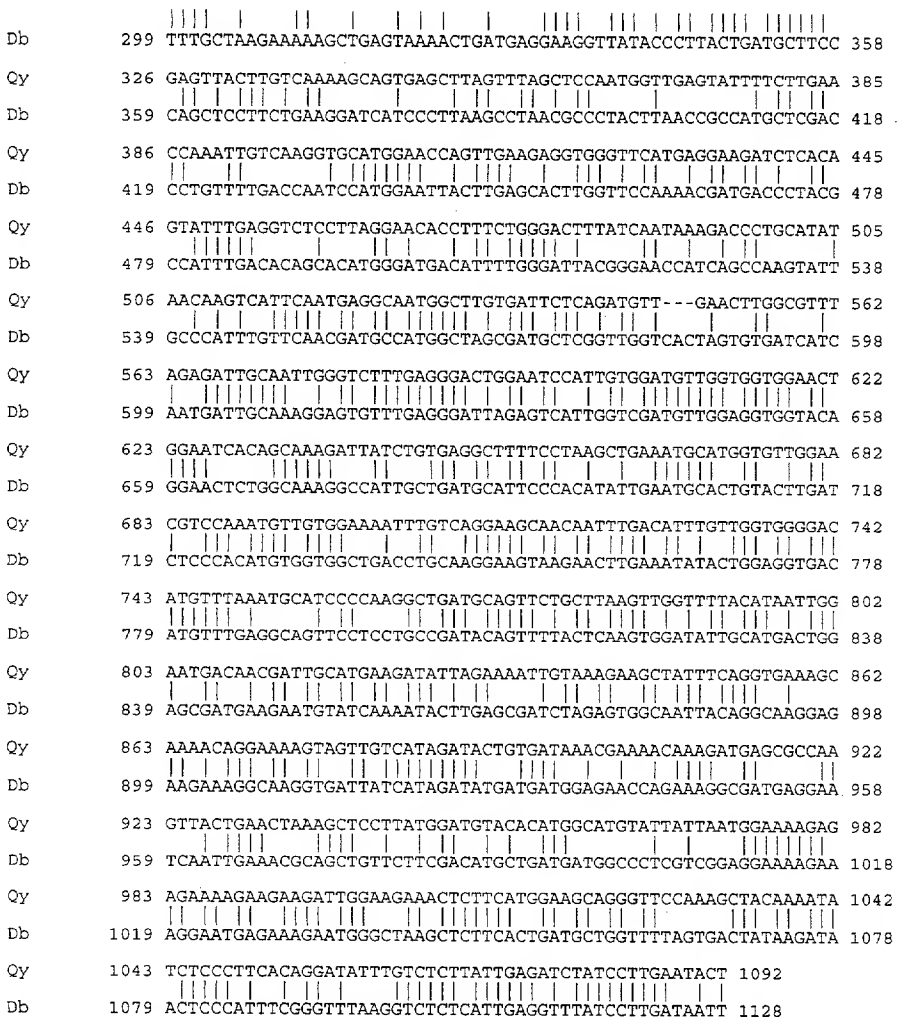
```
Qy         29 AATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTT 88
Db         53 AATGGAGAGCATTCCAACGAGCTACTTCATGCTCAAGCCCACATCTGGAACCACATCTTC 112

Qy         89 GGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATC 148
Db        113 AGCTTCATAAACTCCATGTCCCTCAAATCTGCAATTCAACTAGGTATACCAGATATCATC 172

Qy        149 CACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTC 208
Db        173 AACAAACATGGC---------CCCATGACTCTTTCTGAGCTCACATCTGCACTACCAATC 223

Qy        209 CCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTC 268
Db        224 CATCCAACCAAATCCCACAGCGTCTACCGCCTCATGCGAATATTGGTGCACTCTGGCTTC 283

Qy        269 TTTGAGATAGTAAGAATCCATGACA---ACATAGAAGCATATGCTCTCACTGCTGCTTCA 325
Db        284 TTTGCTAAGAAAAAGCTGAGTAAAACTGATGAGGAAGGTTATACCCTTACTGATGCTTCC 343

Qy        326 GAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAA 385
Db        344 CAGCTCCTTCTGAAGGATCATCCCTTAAGCCTAACGCCCTTCTTAACCGCCATGCTCGAC 403

Qy        386 CCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACA 445
Db        404 CCTGTTTTGACCACACCATGGAATTACTTGAGCACTTGGTTCCAAAACGAGGACCCTACG 463

Qy        446 GTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATAT 505
Db        464 CCATTTGACACAGCACATGGGATGACATTTTGGGATTACGGGAACCATCAGCCAAGTATT 523

Qy        506 AACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTA-- 563
Db        524 GCCCATTTGTTCAATGATGCCATGGCTAGCGATGCTCGGTTGGTCACCAGCGTGATCATC 583

Qy        564 -GAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACT 622
Db        584 GATGATTGCAAAGGAGTGTTTGAGGGATTAGAGTCATTGGTCGATGTTGGAGGTGGTACA 643

Qy        623 GGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAA 682
Db        644 GGAACTGTGGCAAAGGCTATTGCTGATGCATTCCCACATATTGAATGCACTGTACTTGAT 703

Qy        683 CGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGAC 742
Db        704 CTCCCACATGTGGTGGCTGACCTGCAAGGAAGTAAGAACTTGAAATATACTGGAGGTGAC 763

Qy        743 ATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGG 802
Db        764 ATGTTTGAGGCAGTTCCTCCTGCCGATACAGTTTTACTCAAGTGGATATTGCATGACTGG 823

Qy        803 AATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGC 862
Db        824 AACGATGAAGAATGTATCAAAATACTTAAGCGATCTAGAGTGGCAATTACAAGCAAGGAC 883

Qy        863 AAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAA 922
Db        884 AAGAAAGGCAAGGTGATTATCATAGATATGATGATGGAGAACCAGAAGGGGGATGAGGAA 943

Qy        923 GTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAG 982
Db        944 TCAATTGAAACGCAGCTGTTCTTCGACATGCTGATGATGGCCCTCGTCAGAGGACAAGAA 1003

Qy        983 AGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATA 1042
Db       1004 AGGAATGAGAAAGAATGGGCTAAGCTCTTCACTGATGCTGGTTTCAGTGACTATAAGATA 1063

Qy       1043 TCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACT 1092
Db       1064 ACTCCCATTTTGGGTTTAAGGTCTCTCATTGAGGTTTATCCTTGATAATT 1113
```

```
RESULT 13
RCH439742
LOCUS       RCH439742               1265 bp    mRNA    linear   PLN 06-JAN-2003
DEFINITION  Rosa chinensis mRNA for orcinol O-methyltransferase (oomt2 gene).
ACCESSION   AJ439742
VERSION     AJ439742.1  GI:27527923
KEYWORDS    oomt2 gene; orcinol O-methyltransferase.
SOURCE      Rosa chinensis
  ORGANISM  Rosa chinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
REFERENCE   1  (bases 1 to 1265)
  AUTHORS   Cock,J.Mark., Scalliet,G. and Hugueney,P.
  TITLE     Characterisation of a novel O-methyltransferase involved in the
            biosynthesis of 3,5-dimethoxytoluene and 1,3,5-trimethoxybenzene,
            two major scent components of rose flowers
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1265)
  AUTHORS   Cock,J.Mark.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-MAR-2002) RDP, ENSL, 46, allee d'Italie, Lyon cedex
            07 69364, France
FEATURES             Location/Qualifiers
     source          1. .1265
                     /organism="Rosa chinensis"
                     /mol_type="mRNA"
                     /cultivar="Old Blush"
                     /db_xref="taxon:74649"
     gene            1. .1265
                     /gene="oomt2"
     CDS             20. .1120
                     /gene="oomt2"
                     /codon_start=1
                     /product="orcinol O-methyltransferase"
                     /protein_id="CAD29459.1"
```

```
                    /db_xref="GI:27527924"
                    /translation="MERLNSFKHLNQKWSNGEHSNELLHAQAHIWNHIFSFINSMSLK
                    SAIQLGIPDIINKHGPMTLSELTSALPIHPTKSHSVYRLMRILVHSGFFAKKKLSKTD
                    EEGYTLTDASQLLLKDHPLSLTPFLTAMLDPVLTTPWNYLSTWFQNEDPTPFDTAHGM
                    TFWDYGNHQPSIAHLFNDAMASDARLVTSVIIDDCKGVFEGLESLVDVGGGTGTVAKA
                    IADAFPHIECTVLDLPHVVADLQGSKNLKYTGGDMFEAVPPADTVLLKWILHDWNDEE
                    CIKILKRSRVAITSKDKKGKVIIIDMMMENQKGDEESIETQLFFDMLMMALVRGQERN
                    EKEWAKLFTDAGFSDYKITPILGLRSLIEVYP"
BASE COUNT      400 a    273 c    265 g    327 t
ORIGIN

  Query Match             26.1%;  Score 302.8;  DB 8;  Length 1265;
  Best Local Similarity   57.3%;  Pred. No. 2.1e-58;
  Matches  613;  Conservative    0;  Mismatches  442;  Indels   15;  Gaps    3;

Qy         29 AATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTT 88
Db         65 AATGGAGAGCATTCCAACGAGCTACTTCATGCTCAAGCCCACATCTGGAACCACATCTTC 124

Qy         89 GGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATC 148
Db        125 AGCTTCATAAACTCCATGTCCCTCAAATCTGCAATTCAACTAGGTATACCAGATATCATC 184

Qy        149 CACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTC 208
Db        185 AACAAACATGGC---------CCCATGACTCTTTCTGAGCTCACATCTGCACTACCAATC 235

Qy        209 CCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTC 268
Db        236 CATCCAACCAAATCCCACAGCGTCTACCGCCTCATGCGAATATTGGTGCACTCTGGCTTC 295

Qy        269 TTTGAGATAGTAAGAATCCATGACA---ACATAGAAGCATATGCTCTCACTGCTGCTTCA 325
Db        296 TTTGCTAAGAAAAAGCTGAGTAAAACTGATGAGGAAGGTTATACCCTTACTGATGCTTCC 355

Qy        326 GAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAA 385
Db        356 CAGCTCCTTCTGAAGGATCATCCCTTAAGCCTAACGCCCTTCTTAACCGCCATGCTCGAC 415

Qy        386 CCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACA 445
Db        416 CCTGTTTTGACCACACCATGGAATTACTTGAGCACTTGGTTCCAAAACGAGGACCCTACG 475

Qy        446 GTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATAT 505
Db        476 CCATTTGACACAGCACATGGGATGACATTTTGGGATTACGGGAACCATCAGCCAAGTATT 535

Qy        506 AACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTA-- 563
Db        536 GCCCATTTGTTCAATGATGCCATGGCTAGCGATGCTCGGTTGGTCACCAGCGTGATCATC 595

Qy        564 -GAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACT 622
Db        596 GATGATTGCAAAGGAGTGTTTGAGGGATTAGAGTCATTGGTCGATGTTGGAGGTGGTACA 655

Qy        623 GGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAA 682
Db        656 GGAACTGTGGCAAAGGCTATTGCTGATGCATTCCCACATATTGAATGCACTGTACTTGAT 715

Qy        683 CGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGAC 742
Db        716 CTCCCACATGTGGTGGCTGACCTGCAAGGAAGTAAGAACTTGAAATATACTGGAGGTGAC 775

Qy        743 ATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGG 802
Db        776 ATGTTTGAGGCAGTTCCTCCTGCCGATACAGTTTTACTCAAGTGGATATTGCATGACTGG 835

Qy        803 AATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGC 862
Db        836 AACGATGAAGAATGTATCAAAATACTTAAGCGATCTAGAGTGGCAATTACAAGCAAGGAC 895

Qy        863 AAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAA 922
Db        896 AAGAAAGGCAAGGTGATTATCATAGATATGATGATGGAGAACCAGAAGGGGGATGAGGAA 955

Qy        923 GTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAG 982
Db        956 TCAATTGAAACGCAGCTGTTCTTCGACATGCTGATGATGGCCCTCGTCAGAGGACAAGAA 1015

Qy        983 AGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATA 1042
Db       1016 AGGAATGAGAAAGAATGGGCTAAGCTCTTCACTGATGCTGGTTTCAGTGACTATAAGATA 1075

Qy       1043 TCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACT 1092
Db       1076 ACTCCCATTTTGGGTTTAAGGTCTCTCATTGAGGTTTATCCTTGATAATT 1125


RESULT 14
AF502433
LOCUS       AF502433                1252 bp    mRNA    linear   PLN 14-AUG-2002
DEFINITION  Rosa hybrid cultivar orcinol O-methyltransferase (OOMT1) mRNA,
            complete cds.
ACCESSION   AF502433
VERSION     AF502433.1  GI:20514366
KEYWORDS    .
SOURCE      Rosa hybrid cultivar
  ORGANISM  Rosa hybrid cultivar
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
REFERENCE   1  (bases 1 to 1252)
  AUTHORS   Lavid,N., Wang,J., Shalit,M., Guterman,I., Bar,E., Beuerle,T.,
            Menda,N., Shafir,S., Zamir,D., Adam,Z., Vainstein,A., Weiss,D.,
            Pichersky,E. and Lewinsohn,E.
  TITLE     O-methyltransferases involved in the biosynthesis of volatile
            phenolic derivatives in rose petals
  JOURNAL   Plant Physiol. 129 (4), 1899-1907 (2002)
  MEDLINE   22167322
   PUBMED   12177504
REFERENCE   2  (bases 1 to 1252)
  AUTHORS   Lavid,N., Wang,J., Shalit,M., Guterman,I., Bar,E., Beuerle,T.,
            Menda,N., Shafir,S., Zamir,D., Adam,Z., Vainstein,A., Weiss,D.,
            Pichersky,E. and Lewinsohn,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-2002) Vegetables Crops, Newe Ya'ar Research
            Center, Agricultural Research Organization, P.O. Box 1021, Ramat
            Yishay 30095, Israel
FEATURES             Location/Qualifiers
     source          1..1252
                     /organism="Rosa hybrid cultivar"
                     /mol_type="mRNA"
                     /cultivar="Fragrant Cloud"
                     /isolate="fc0496"
                     /db_xref="taxon:128735"
                     /tissue_type="petals"
     gene            1..1252
                     /gene="OOMT1"
     CDS             11..1114
                     /gene="OOMT1"
                     /function="catalyzes the S-adenosylmethionine-dependent
                     O-methylation of orcinol (3,5-dihydroxytoluene) to form
                     3-methoxy, 5 hydroxytoluene"
                     /codon_start=1
                     /product="orcinol O-methyltransferase"
                     /protein_id="AAM23004.1"
                     /db_xref="GI:20514367"
                     /translation="MERLNSFRHLNQKWSNGEHSNELLHAQAHIWNHIFSFINSMSLK
                     SAIQLGIPDIINKHGYPMTLSELTSALPIHPTKSHSVYRLMRILVHSGFFAKKKLSKT
                     DEEGYTLTDASQLLLKDHPLSLTPYLTAMLDPVLTNPWNYLSTWFQNDDPTPFDTAHG
                     MTFWDYGNHQPSIAHLFNDAMASDARLVTSVIINDCKGVFEGLESLVDVGGGTGTLAK
                     AIADAFPHIECTVLDLPHVVADLQGSKNLKYTGGDMFEAVPPADTVLLKWILHDWSDE
                     ECIKILERSRVAITGKEKKGKVIIIDMMMENQKGDEESIETQLFFDMLMMALVGGKER
                     NEKEWAKLFTDAGFSDYKITPISGLRSLIEVYP"
BASE COUNT      383 a    270 c    266 g    333 t
ORIGIN
```

```
Query Match             26.1%;  Score 302.4;  DB 8;  Length 1252;
Best Local Similarity   57.2%;  Pred. No. 2.6e-58;
Matches  612;  Conservative    0;  Mismatches  446;  Indels   12;  Gaps    3;

Qy         29 AATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTT 88
Db         56 AATGGAGAGCATTCCAACGAGCTACTTCATGCTCAAGCCCACATCTGGAACCACATCTTC 115

Qy         89 GGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATC 148
Db        116 AGCTTCATAAACTCCATGTCCCTCAAATCTGCAATTCAACTAGGTATACCAGATATCATC 175

Qy        149 CACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTC 208
Db        176 AACAAA------CATGGTTACCCCATGACTCTTTCTGAGCTCACATCTGCACTACCAATC 229

Qy        209 CCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTC 268
Db        230 CATCCAACCAAATCCCACAGCGTCTACCGCCTCATGCGAATATTGGTGCACTCTGGCTTC 289

Qy        269 TTTGAGATAGTAAGAATCCATGACA---ACATAGAAGCATATGCTCTCACTGCTGCTTCA 325
Db        290 TTTGCTAAGAAAAAGCTGAGTAAAACTGATGAGGAAGGTTATACCCTTACTGATGCTTCC 349

Qy        326 GAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAA 385
Db        350 CAGCTCCTTCTGAAGGATCATCCCTTAAGCCTAACGCCCTACTTAACCGCCATGCTCGAC 409

Qy        386 CCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACA 445
Db        410 CCTGTTTTGACCAATCCATGGAATTACTTGAGCACTTGGTTCCAAAACGATGACCCTACG 469

Qy        446 GTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATAT 505
Db        470 CCATTTGACACAGCACATGGGATGACATTTTGGGATTACGGGAACCATCAGCCAAGTATT 529

Qy        506 AACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTT---GAACTTGGCGTTT 562
Db        530 GCCCATTTGTTCAACGATGCCATGGCTAGCGATGCTCGGTTGGTCACTAGTGTGATCATC 589

Qy        563 AGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACT 622
Db        590 AATGATTGCAAAGGAGTGTTTGAGGGATTAGAGTCATTGGTCGATGTTGGAGGTGGTACA 649

Qy        623 GGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAA 682
Db        650 GGAACTCTGGCAAAGGCCATTGCTGATGCATTCCCACATATTGAATGCACTGTACTTGAT 709

Qy        683 CGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGAC 742
Db        710 CTCCCACATGTGGTGGCTGACCTGCAAGGAAGTAAGAACTTGAAATATACTGGAGGTGAC 769

Qy        743 ATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGG 802
Db        770 ATGTTTGAGGCAGTTCCTCCTGCCGATACAGTTTTACTCAAGTGGATATTGCATGACTGG 829

Qy        803 AATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGC 862
Db        830 AGCGATGAAGAATGTATCAAAATACTTGAGCGATCTAGAGTGGCAATTACAGGCAAGGAG 889

Qy        863 AAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAA 922
Db        890 AAGAAAGGCAAGGTGATTATCATAGATATGATGATGGAGAACCAGAAAGGCGATGAGGAA 949

Qy        923 GTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAG 982
Db        950 TCAATTGAAACGCAGCTGTTCTTCGACATGCTGATGATGGCCCTCGTCGGAGGAAAAGAA 1009

Qy        983 AGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATA 1042
Db       1010 AGGAATGAGAAAGAATGGGCTAAGCTCTTCACTGATGCTGGTTTTAGTGACTATAAGATA 1069
```

```
Qy       1043 TCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACT 1092
Db       1070 ACTCCCATTTCGGGTTTAAGGTCTCTCATTGAGGTTTATCCTTGATAATT 1119


RESULT 15
RCH439741
LOCUS       RCH439741               1275 bp    mRNA    linear   PLN 06-JAN-2003
DEFINITION  Rosa chinensis mRNA for orcinol O-methyltransferase (oomt1 gene).
ACCESSION   AJ439741
VERSION     AJ439741.1  GI:27527921
KEYWORDS    oomt1 gene; orcinol O-methyltransferase.
SOURCE      Rosa chinensis
  ORGANISM  Rosa chinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
REFERENCE   1  (bases 1 to 1275)
  AUTHORS   Cock,J.Mark., Scalliet,G. and Hugueney,P.
  TITLE     Characterisation of a novel O-methyltransferase involved in the
            biosynthesis of 3,5-dimethoxytoluene and 1,3,5-trimethoxybenzene,
            two major scent components of rose flowers
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1275)
  AUTHORS   Cock,J.Mark.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-MAR-2002) RDP, ENSL, 46, allee d'Italie, Lyon cedex
            07 69364, France
FEATURES             Location/Qualifiers
     source          1. .1275
                     /organism="Rosa chinensis"
                     /mol_type="mRNA"
                     /cultivar="Old Blush"
                     /db_xref="taxon:74649"
     gene            1. .1275
                     /gene="oomt1"
     CDS             20. .1123
                     /gene="oomt1"
                     /codon_start=1
                     /product="orcinol O-methyltransferase"
                     /protein_id="CAD29458.1"
                     /db_xref="GI:27527922"
                     /translation="MERLNSFRHLNQKWSNGEHSNELLHAQAHIWNHIFSFINSMSLK
                     SAIQLGIPDIINKHGYPMTLSELTSALPIHPTKSHSVYRLMRILVHSGFFAKKKLSKT
                     DEEGYTLTDASQLLLKDHPLSLTPYLTAMLDPVLTNPWNYLSTWFQNDDPTPFDTAHG
                     MTFWDYGNHQPSIAHLFNDAMASDARLVTSVIINDCKGVFEGLESLVDVGGGTGTLAK
                     AIADAFPHIECTVLDLPHVVADLQGSKNLKYTGGDMFEAVPPADTVLLKWILHDWSDE
                     ECIKILERSRVAITGKEKKGKVIIIDMMMENQKGDEESIETQLFFDMLMMALVGGKER
                     NEKEWAKLFTDAGFSDYKITPISGLRSLIEVYP"
BASE COUNT      402 a    273 c    266 g    334 t
ORIGIN

  Query Match             26.1%;  Score 302.4;  DB 8;  Length 1275;
  Best Local Similarity   57.2%;  Pred. No. 2.5e-58;
  Matches  612;  Conservative    0;  Mismatches  446;  Indels   12;  Gaps    3;

Qy         29 AATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTT 88
Db         65 AATGGAGAGCATTCCAACGAGCTACTTCATGCTCAAGCCCACATCTGGAACCACATCTTC 124

Qy         89 GGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATC 148
Db        125 AGCTTCATAAACTCCATGTCCCTCAAATCTGCAATTCAACTAGGTATACCAGATATCATC 184

Qy        149 CACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTC 208
Db        185 AACAAA------CATGGTTACCCCATGACTCTTTCTGAGCTCACATCTGCACTACCAATC 238

Qy        209 CCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTC 268
Db        239 CATCCAACCAAATCCCACAGCGTCTACCGCCTCATGCGAATATTGGTGCACTCTGGCTTC 298

Qy        269 TTTGAGATAGTAAGAATCCATGACA---ACATAGAAGCATATGCTCTCACTGCTGCTTCA 325
```



Search completed: November 10, 2003, 01:02:44
Job time : 4628 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 9, 2003, 23:29:44 ; Search time 428 Seconds
(without alignments)
7316.234 Million cell updates/sec

Title: US-09-868-547-3
Perfect score: 1160
Sequence: 1 gtttgctattatggcttctt.....ataaaagttatctttgtgtt 1160

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1160	100.0	1160	21	AAA29322	Glycine max isofla
2	645.4	55.6	1231	22	AAC89152	Alfalfa isoflavone
3	508.6	43.8	1253	21	AAA29324	Glycine max isofla
4	507.2	43.7	1065	21	AAA29323	Glycine max isofla
5	507.2	43.7	1065	21	AAA29325	Glycine max isofla
6	497.8	42.9	1053	21	AAA29321	Glycine max isofla
7	209.2	18.0	1041	20	AAX90136	Coptis japonica no
8	178.4	15.4	1050	20	AAX89255	(S)-3'-hydroxy-N-m
9	158.2	13.6	256	25	ABX31899	Human GDP-mannose
10	142.8	12.3	266	25	ABX23518	Human GDP-mannose
11	142.2	12.3	381	25	ABX19667	Human GDP-mannose
12	120.2	10.4	1412	20	AAX25200	Maize caffeic O-me
13	118	10.2	1098	22	AAD19546	Medicago sativa ca
14	118	10.2	1341	15	AAQ77692	Alfalfa COMT clone
15	114	9.8	1267	21	AAC40707	Arabidopsis thalia
16	114	9.8	1314	24	AAD27047	Soybean caffeic ac
17	112.4	9.7	1306	21	AAC49059	Arabidopsis thalia
18	112.4	9.7	1308	21	AAC40838	Arabidopsis thalia
19	111.8	9.6	377	25	ABX21443	Human GDP-mannose
20	111.6	9.6	1315	20	AAX25201	Maize caffeic O-me
21	109.8	9.5	1368	14	AAQ38448	pPLC4 coding seque
22	109.8	9.5	1501	24	ABK48067	cDNA encoding 5-hy
23	109.8	9.5	1503	20	AAX26295	Aspen bispecific O
24	109	9.4	279	25	ABX28928	Human GDP-mannose
25	109	9.4	1507	24	ABS54111	Tobacco caffeic ac
26	108	9.3	1244	14	AAQ38449	pOMT1.A and pOMT1.
27	108	9.3	1460	20	AAX35662	DNA encoding a tob
28	107.8	9.3	1370	24	AAD27040	Soybean caffeic ac
29	107.6	9.3	272	25	ABX31152	Human GDP-mannose
30	107.4	9.3	1418	20	AAX35663	DNA encoding a tob
31	107.4	9.3	1431	14	AAQ38450	OMTI coding sequen
32	106.8	9.2	1306	20	AAX25202	Maize caffeic O-me
33	106.2	9.2	978	21	AAC42978	Arabidopsis thalia
34	100.2	8.6	1149	24	ABZ14777	Arabidopsis thalia
35	100.2	8.6	1283	21	AAC46334	Arabidopsis thalia
36	99.4	8.6	1286	21	AAC41985	Arabidopsis thalia
37	95.4	8.2	1195	24	AAD27044	Soybean caffeic ac
38	94.6	8.2	1380	20	AAX81875	DNA encoding a bif
39	94.6	8.2	1380	21	AAA88683	Sweetgum bifunctio
40	94.6	8.2	1380	22	AAD11127	L. styraciflua ang
41	92.6	8.0	1630	21	AAA68014	Eucalyptus grandis
42	90.8	7.8	1486	19	AAV71042	cDNA encoding (iso
43	90.8	7.8	1486	25	AAD47178	Clarkia breweri IE
44	89.6	7.7	381	21	AAA68081	Eucalyptus grandis
45	88.8	7.7	1201	21	AAZ88734	P. sylvestris PMT

ALIGNMENTS

RESULT 1
AAA29322
ID AAA29322 standard; cDNA; 1160 BP.
XX
AC AAA29322;
XX
DT 26-SEP-2000 (first entry)
XX
DE Glycine max isoflavone O-methyltransferase clone sr1.pk0015.b4.
XX
KW Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW Flavonoid; flower colour; growth; pollination; irradiation; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 11..1087
FT /*tag= a
FT /product= Isoflavone_O-methyltransferase
XX
PN WO200037656-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US30338.
XX
PR 21-DEC-1998; 98US-0113190.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Famodu OO, McGonigle B, Odell JT, Fader GM, Palco SC;

XX
DR WPI; 2000-442680/38.
DR P-PSDB; AAY96584.
XX
PT New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT transgenic plants and for immunological screening of cDNA libraries
XX
PS Claim 2; Page 32; 39pp; English.
XX
CC AAA29321-25 encode isoflavone O-methyltransferases isolated from various
CC soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c and ssl1c
CC were prepared from soybean embryo (19 days after flowering), root, 8-day
CC old root and seed (25 days after flowering).
CC Isoflavone O-methyltransferase catalyses the first step in degradation
CC of daidzein. Suppression of this enzyme will yield higher concentrations
CC of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC co-pigments in flower colour, stimulate pollen tube growth, attract
CC pollenators, act as feeding deterrents and protect against UV
CC irradiation in fruits and seeds. The cDNA and proteins can be used to
CC isolate homologues, for immunological screening and for positive
CC selection methods.
XX
SQ Sequence 1160 BP; 361 A; 197 C; 254 G; 348 T; 0 other;

Query Match 100.0%; Score 1160; DB 21; Length 1160;
Best Local Similarity 100.0%; Pred. No. 3.4e-295;
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

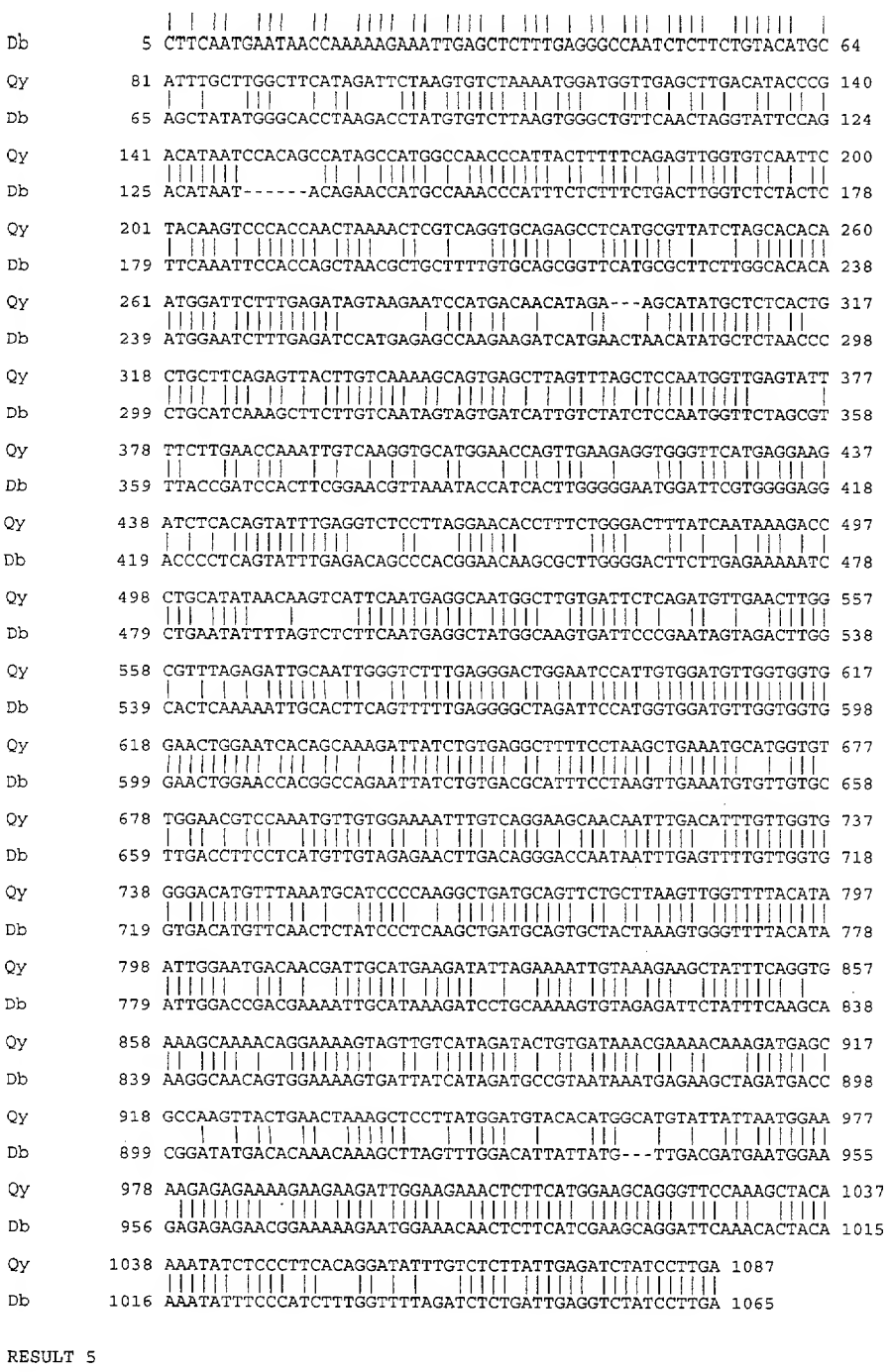
```
Qy          1 GTTTGCTATTATGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTTGCTATTATGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGG 60

Qy         61 TCAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGAT 120

Qy        121 GGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTT 180

Qy        181 TTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCT 240

Qy        241 CATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGA 300

Qy        301 AGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGC 360

Qy        361 TCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAG 420

Qy        421 GTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGA 480

Qy        481 CTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTC 540

Qy        541 TCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCAT 600

Qy        601 TGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAA 660

Qy        661 GCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAA 720

Qy        721 TTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCT 780

Qy        781 TAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAA 840

Qy        841 AGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAA 900

Qy        901 CGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGC 960

Qy        961 ATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGC 1020

Qy       1021 AGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTA 1080

Qy       1081 TCCTTGAATACTGACGCTGCAATATTCCATTTAGTAGTTAATTTGCATGTTATCAATAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCCTTGAATACTGACGCTGCAATATTCCATTTAGTAGTTAATTTGCATGTTATCAATAAA 1140

Qy       1141 ATAAAAGTTATCTTTGTGTT 1160
              ||||||||||||||||||||
Db       1141 ATAAAAGTTATCTTTGTGTT 1160
```

RESULT 2
AAC89152
ID AAC89152 standard; cDNA; 1231 BP.
XX
AC AAC89152;
XX
DT 07-MAR-2001 (first entry)
XX
DE Alfalfa isoflavone 4'-O-methyl transferase 8 coding sequence.
XX
KW Alfalfa; isoflavone 4'-O-methyl transferase 8; IOMT; transgenic plant;
KW disease resistance; ss.
XX
OS Medicago sativa.
XX
PN WO200071736-A1.
XX
PD 30-NOV-2000.
XX
PF 15-MAY-2000; 2000WO-US13389.
XX
PR 20-MAY-1999; 99US-0135026.
XX
PA (ROBE-) ROBERTS NOBLE FOUND INC SAMMUEL.
XX
PI Dixon RA, He XZ;
XX
DR WPI; 2001-061355/07.
XX
PT Genetically manipulating biologically active 4'-O-methylated
PT isoflavonoid for producing plants with increased disease resistance,
PT and for accumulation of 4'-O-methylated isoflavonoid nutraceuticals in
PT plants -
XX
PS Claim 5; Fig 2; 49pp; English.
XX
CC The present sequence is isoflavone 4'-O-methyl transferase 8 (IOMT)

```
CC   coding sequence from Alfalfa. The present invention relates to a method
CC   for genetically manipulating the level of biologically active IOMT. The
CC   method involves transforming the target plant with a DNA fragment (the
CC   present sequence) to form a transgenic plant. IOMT is useful as
CC   foodstuff, nutritional supplement, animal feed supplement, nutraceutical
CC   and as a pharmaceutical. Transgenic expression of IOMT in legumes is
CC   useful for engineering both phytoalexin levels for improved disease
CC   resistance, and health promoting nutraceutical phytochemicals. IOMT is
CC   useful for engineering isoflavone 4'-O-methylation, or other organisms
CC   that do not naturally produce isoflavonoids.
XX
SQ   Sequence 1231 BP; 402 A; 183 C; 235 G; 411 T; 0 other;

  Query Match             55.6%;  Score 645.4;  DB 22;  Length 1231;
  Best Local Similarity   76.0%;  Pred. No. 9.7e-160;
  Matches  841;  Conservative    0;  Mismatches  251;  Indels   15;  Gaps    3;

Qy         18 CTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACA 77
Db         29 CTTCATCAATTAATGGCCGAAAACCAAGTGAAATTTTCAAAGCACAAGCTTTATTATACA 88

Qy         78 AACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATAC 137
Db         89 AACATATATATGCCTTCATAGATTCCATGTCTCTTAAATGGGCTGTTGAAATGAACATAC 148

Qy        138 CCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAA 197
Db        149 CAAACATAAT------CCAAAACCATGGCAAACCAATTTCTCTTTCAAACTTAGTTTCAA 202

Qy        198 TTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCAC 257
Db        203 TTCTTCAAGTTCCATCGTCGAAAATAGGTAACGTGCGGCGTCTCATGCGTTACCTCGCGC 262

Qy        258 ACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTG 317
Db        263 ACAACGGATTCTTCGAGATAATTACAAAAGAAGA------AGAGTCTTATGCTCTCACTG 316

Qy        318 CTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATT 377
Db        317 TTGCTTCAGAGCTTCTTGTTAGAGGCAGTGATCTTTGTTTAGCACCAATGGTTGAGTGTG 376

Qy        378 TTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAG 437
Db        377 TTCTTGATCCAACTCTTTCGGGTTCGTATCATGAGCTGAAGAAATGGATTTATGAGGAAG 436

Qy        438 ATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACC 497
Db        437 ATCTTACACTCTTTGGTGTTACTTTAGGATCTGGTTTTTGGGATTTTCTTGATAAAAATC 496

Qy        498 CTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGG 557
Db        497 CTGAATATAATACCTCATTTAATGATGCAATGGCTAGTGATTCTAAATTGATAAACTTGG 556

Qy        558 CGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTG 617
Db        557 CATTGAGAGATTGTGATTTTGTGTTTGATGGATTGGAATCAATTGTGGATGTTGGTGGTG 616

Qy        618 GAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGT 677
Db        617 GAACTGGAACAACTGCTAAGATTATTTGTGAGACTTTTCCTAAGTTGAAATGTATTGTGT 676

Qy        678 TGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTG 737
Db        677 TTGATAGGCCACAAGTTGTAGAGAACTTATCTGGAAGCAATAATTTGACTTATGTTGGTG 736

Qy        738 GGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATA 797
Db        737 GGGACATGTTCACATCTATTCCTAATGCTGATGCAGTTTTGCTTAAGTATATTCTACATA 796

Qy        798 ATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTG 857
Db        797 ATTGGACTGATAAGGATTGCCTAAGGATACTGAAGAAATGTAAAGAAGCTGTTACAAATG 856

Qy        858 AAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGC 917
Db        857 ATGGGAAAAGAGGAAAAGTGACTATTATAGACATGGTGATAGATAAAAAAAAAGATGAGA 916

Qy        918 GCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAA 977
Db        917 ATCAAGTTACTCAAATTAAGCTCCTTATGGATGTAAACATGGCTTG---TCTAAATGGAA 973

Qy        978 AAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACA 1037
Db        974 AAGAGAGAAATGAGGAAGAATGGAAGAAACTCTTCATAGAAGCTGGTTTCCAACACTATA 1033

Qy       1038 AAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGC 1097
Db       1034 AGATATCTCCTTTGACTGGATTTTTGTCTCTTATTGAGATCTATCCATAAACACTTTTGC 1093

Qy       1098 TGCAATATTCCATTTAGTAGTTAATTT 1124
Db       1094 TTTGATCATTCATCCATTCTATTGTTT 1120


RESULT 3
AAA29324
ID   AAA29324 standard; cDNA; 1253 BP.
XX
AC   AAA29324;
XX
DT   26-SEP-2000  (first entry)
XX
DE   Glycine max isoflavone O-methyltransferase clone ssl1c.pk002.d1.
XX
KW   Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW   Flavonoid; flower colour; growth; pollination; irradiation; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..1048
FT                   /*tag=  a
FT                   /product= Isoflavone_O-Methyltransferase
FT                   /partial
XX
PN   WO200037656-A2.
XX
PD   29-JUN-2000.
XX
PF   20-DEC-1999;   99WO-US30338.
XX
PR   21-DEC-1998;   98US-0113190.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Famodu OO,  McGonigle B,  Odell JT,  Fader GM,  Falco SC;
XX
DR   WPI; 2000-442680/38.
DR   P-PSDB; AAY96586.
XX
PT   New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT   transgenic plants and for immunological screening of cDNA libraries
XX
PS   Claim 2; Page 35; 39pp; English.
XX
CC   AAA29321-25 encode isoflavone O-methyltransferases isolated from various
CC   soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c and ssl1c
CC   were prepared from soybean embryo (19 days after flowering), root, 8-day
CC   old root and seed (25 days after flowering).
CC   Isoflavone O-methyltransferase catalyses the first step in degradation
CC   of daidzein. Suppression of this enzyme will yield higher concentrations
CC   of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC   co-pigments in flower colour, stimulate pollen tube growth, attract
CC   pollenators, act as feeding deterrents and protect against UV
CC   irradiation in fruits and seeds. The cDNA and proteins can be used to
CC   isolate homologues, for immunological screening and for positive
```

```
CC   selection methods.
XX
SQ   Sequence 1253 BP; 374 A; 232 C; 267 G; 380 T; 0 other;

  Query Match            43.8%;  Score 508.6;  DB 21;  Length 1253;
  Best Local Similarity  68.2%;  Pred. No. 1e-123;
  Matches  754;  Conservative    0;  Mismatches  339;  Indels   12;  Gaps    3;

Qy         34 CCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTTGGCTT 93
Db          1 CCAAAAAGCAATGGAACTCTTTGAGGGCCAATCTCTTTTGTACATGCAGCTATATGGGCA 60

Qy         94 CATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATCCACAG 153
Db         61 CCTAAGACCTATGTGTCTTAAGTGGGCTGTTCAACTAGGTATTCCAGACATAAT------ 114

Qy        154 CCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTCCCACC 213
Db        115 ACAGAACCATGCCAAACCCATTACTCTTTCTGATTTGGTCTCTACTCTTCAAATTCCACC 174

Qy        214 AACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTCTTTGA 273
Db        175 ATCTAAGGCTGGTTTTGTGCAGCAGTTCATGCGCTTTTTGGCACACGATGGAATCTTTGA 234

Qy        274 GAT---AGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTT 330
Db        235 TATCCGTGAGAGCCAAGATGATCATGAATTAGCATATGCTCTAACCCCTGCTTCAAAGCT 294

Qy        331 ACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAA 390
Db        295 TCTAGTTAGTTGCAGTGACCACTGTTTATCTCCAATGGTTCGGATGAATACTGATCCACT 354

Qy        391 TTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATT 450
Db        355 TCTGATGACTACATACCATCACTTTGGGGAATGGATTCGTGGGGAAGACCCCACAGTACA 414

Qy        451 TGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAA 510
Db        415 TGAGACAGCCTTCGGAACAAGCTTTTGGGGACTTCTTGAGAAAAACCCTACACAAATGAG 474

Qy        511 GTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTG 570
Db        475 TCTCTTCAATGAGGCTATGGCAAGTGATTCCCGAATGGTAGACTTGGCACTCAAAAATTG 534

Qy        571 CAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCAC 630
Db        535 CACTTCAGTTTTTGAAGGGCTAGATTCCATGGTGGATGTTGGTGGTGGAACTGGAACCAC 594

Qy        631 AGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAA 690
Db        595 AGCCAGAATTATCTGTGAGGCATTTCCGAAGTTGAAATGTGTTGTGCTTGACCTTCCTCA 654

Qy        691 TGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAA 750
Db        655 TGTTGTAGAAAACTTGACAGGAACCAATAATTTGAGTTTTGTCGGTGGTGACATGTTCAA 714

Qy        751 ATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAA 810
Db        715 CTCTTTCCCTCAAACTGATGCAGTTCTACTAAAGTGGGTTTTACATAATTGGAATGACGA 774

Qy        811 CGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGG 870
Db        775 AAATTGCATAAAGATCCTGAAAAAGTGTAAAGATTCTATTTCAAGCAAAGGCAACAGAGG 834

Qy        871 AAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGA 930
Db        835 AAAAGTGATTATCATAGATATAATAATAAATGAGAAGCTAGATGACCCGGATATGACTCG 894

Qy        931 ACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGA 990
Db        895 AACAAAGCTTAGTTTGGATATAGTTATG---TCGACTATGAATGGAAGAGAGCGAAGTGA 951

Qy        991 AGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTT 1050
```

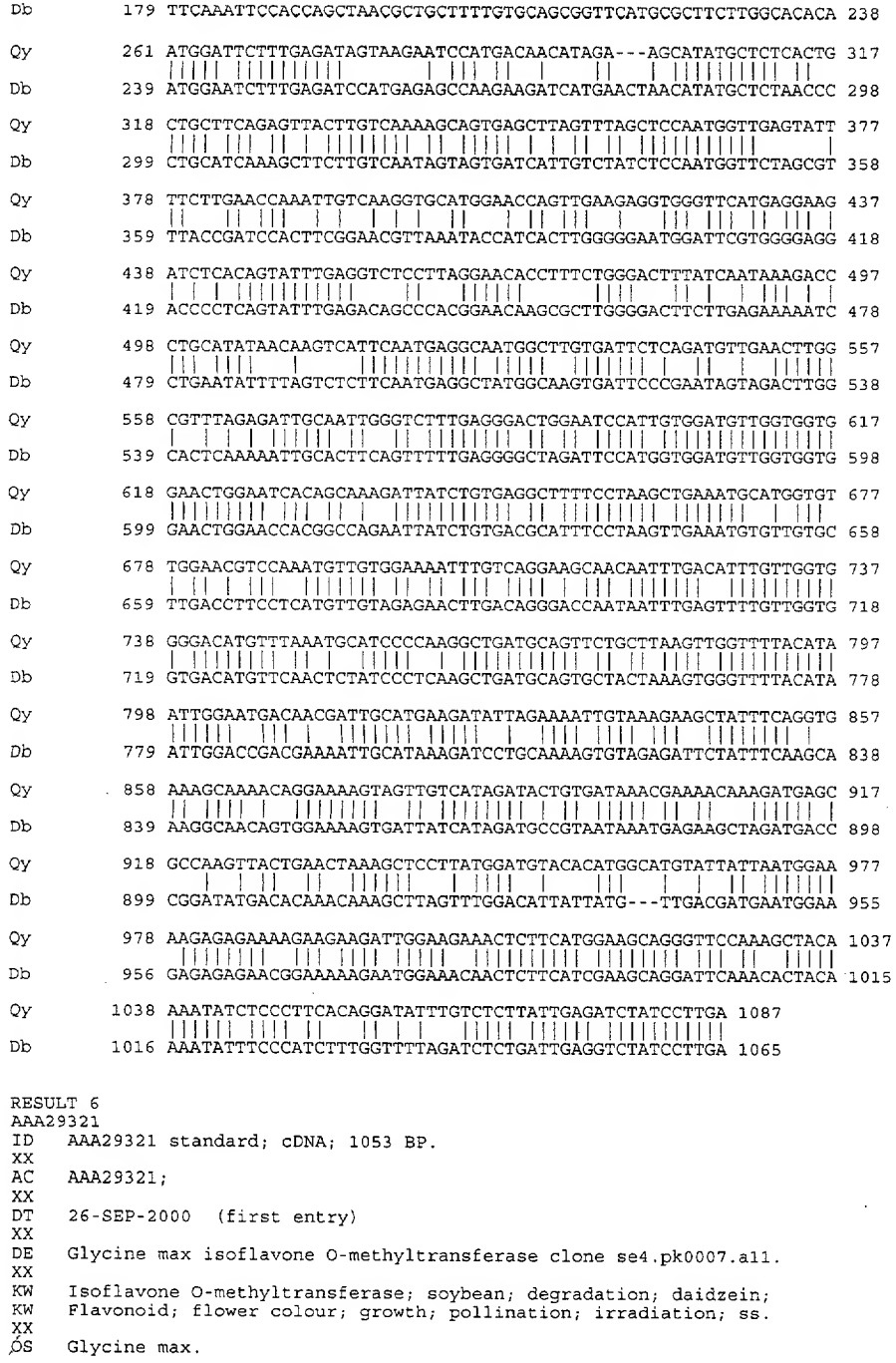
```
Db        952 AAAAGAATGGAAACAAATGTTCATTGAAGCGGGATTCCAACACTGCAAAATATTTCCCAT 1011

Qy       1051 CACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGCTGCAATATTCCAT 1110
Db       1012 CTTTGGTTTCAGATCTCTAATTGAGCTCTATCCTTAGACATATATCCTAGCTTGTTATGC 1071

Qy       1111 TTAGTAGTTAATTTGCATGTTATCA 1135
Db       1072 ATTTTATAGTTTGCGCATGTTGTGA 1096


RESULT 4
AAA29323
ID   AAA29323 standard; cDNA; 1065 BP.
XX
AC   AAA29323;
XX
DT   26-SEP-2000  (first entry)
XX
DE   Glycine max isoflavone O-methyltransferase clone srr1c.pk001.b16.
XX
KW   Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW   Flavonoid; flower colour; growth; pollination; irradiation; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1065
FT                   /*tag=  a
FT                   /transl_except= (pos:268..279, aa:Glu)
FT                   /note= "A 9 bp insertion is present"
XX
PN   WO200037656-A2.
XX
PD   29-JUN-2000.
XX
PF   20-DEC-1999;   99WO-US30338.
XX
PR   21-DEC-1998;   98US-0113190.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Famodu OO,  McGonigle B,  Odell JT,  Fader GM,  Falco SC;
XX
DR   WPI; 2000-442680/38.
DR   P-PSDB; AAY96585.
XX
PT   New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT   transgenic plants and for immunological screening of cDNA libraries
XX
PS   Claim 2; Page 33-34; 39pp; English.
XX
CC   AAA29321-25 encode isoflavone O-methyltransferases isolated from various
CC   soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c and ssl1c
CC   were prepared from soybean embryo (19 days after flowering), root, 8-day
CC   old root and seed (25 days after flowering).
CC   Isoflavone O-methyltransferase catalyses the first step in degradation
CC   of daidzein. Suppression of this enzyme will yield higher concentrations
CC   of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC   co-pigments in flower colour, stimulate pollen tube growth, attract
CC   pollenators, act as feeding deterrents and protect against UV
CC   irradiation in fruits and seeds. The cDNA and proteins can be used to
CC   isolate homologues, for immunological screening and for positive
CC   selection methods.
XX
SQ   Sequence 1065 BP; 315 A; 212 C; 233 G; 305 T; 0 other;

  Query Match            43.7%;  Score 507.2;  DB 21;  Length 1065;
  Best Local Similarity  69.2%;  Pred. No. 2.2e-123;
  Matches  740;  Conservative    0;  Mismatches  318;  Indels   12;  Gaps    3;

Qy         21 CATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAAC 80
```



RESULT 5

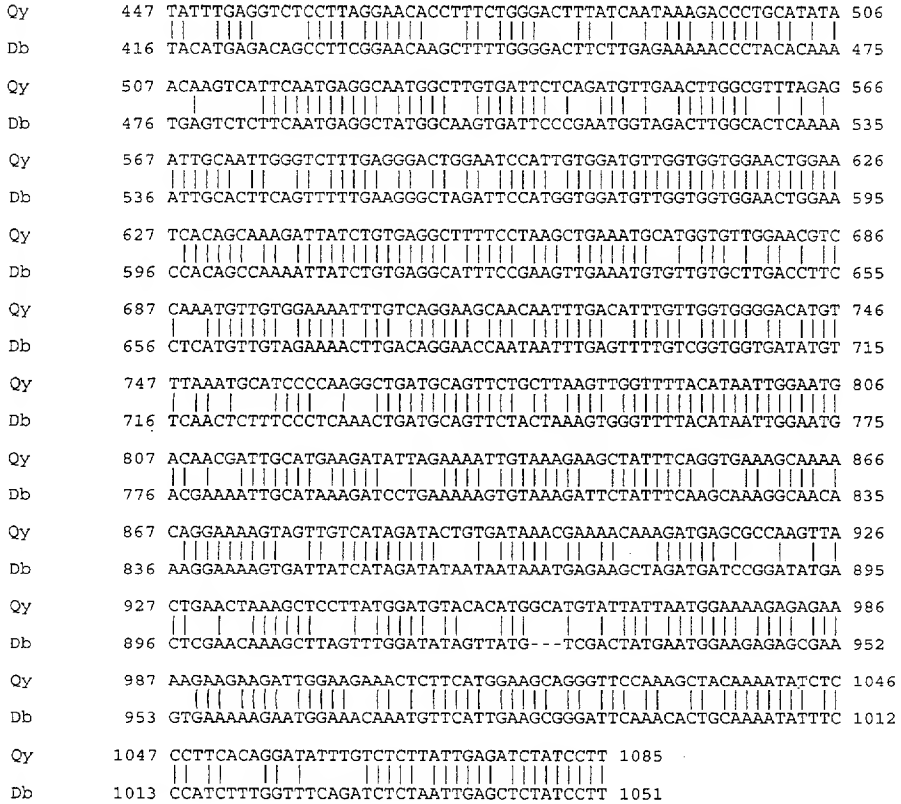
AAA29325
ID AAA29325 standard; cDNA; 1065 BP.
XX
AC AAA29325;
XX
DT 26-SEP-2000 (first entry)
XX
DE Glycine max isoflavone O-methyltransferase clone srr1c.pk001.b16.
XX
KW Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW Flavonoid; flower colour; growth; pollination; irradiation; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 4..1065
FT /*tag= a
FT /transl_except= (pos:268..279, aa:Glu)
FT /note= "A 9 bp insertion is present"
XX
PN WO200037656-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US30338.
XX
PR 21-DEC-1998; 98US-0113190.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;
XX
DR WPI; 2000-442680/38.
DR P-PSDB; AAY96587.
XX
PT New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT transgenic plants and for immunological screening of cDNA libraries
XX
PS Claim 2; Page 37; 39pp; English.
XX
CC AAA29321-25 encode isoflavone O-methyltransferases isolated from various
CC soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c and ssl1c
CC were prepared from soybean embryo (19 days after flowering), root, 8-day
CC old root and seed (25 days after flowering).
CC Isoflavone O-methyltransferase catalyses the first step in degradation
CC of daidzein. Suppression of this enzyme will yield higher concentrations
CC of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC co-pigments in flower colour, stimulate pollen tube growth, attract
CC pollenators, act as feeding deterrents and protect against UV
CC irradiation in fruits and seeds. The cDNA and proteins can be used to
CC isolate homologues, for immunological screening and for positive
CC selection methods.
XX
SQ Sequence 1065 BP; 315 A; 212 C; 233 G; 305 T; 0 other;

Query Match 43.7%; Score 507.2; DB 21; Length 1065;
Best Local Similarity 69.2%; Pred. No. 2.2e-123;
Matches 740; Conservative 0; Mismatches 318; Indels 12; Gaps 3;



RESULT 6
AAA29321
ID AAA29321 standard; cDNA; 1053 BP.
XX
AC AAA29321;
XX
DT 26-SEP-2000 (first entry)
XX
DE Glycine max isoflavone O-methyltransferase clone se4.pk0007.a11.
XX
KW Isoflavone O-methyltransferase; soybean; degradation; daidzein;
KW Flavonoid; flower colour; growth; pollination; irradiation; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 1..1053
FT /*tag= a
FT /product= Isoflavone_O-Methyltransferase
FT /partial
XX
PN WO200037656-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US30338.
XX
PR 21-DEC-1998; 98US-0113190.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;
XX
DR WPI; 2000-442680/38.
DR P-PSDB; AAY96583.
XX
PT New polynucleotide encoding flavonoid biosynthetic enzymes for creating
PT transgenic plants and for immunological screening of cDNA libraries
XX
PS Claim 2; Page 30; 39pp; English.
XX
CC AAA29321-25 encode isoflavone O-methyltransferases isolated from various
CC soybean (Glycine max) tissues. cDNA libraries se4, sr1, srr1c and ssl1c
CC were prepared from soybean embryo (19 days after flowering), root, 8-day
CC old root and seed (25 days after flowering).
CC Isoflavone O-methyltransferase catalyses the first step in degradation
CC of daidzein. Suppression of this enzyme will yield higher concentrations
CC of this beneficial isoflavone in, e.g. soybean seed. Flavonoids serve as
CC co-pigments in flower colour, stimulate pollen tube growth, attract
CC pollenators, act as feeding deterrents and protect against UV
CC irradiation in fruits and seeds. The cDNA and proteins can be used to
CC isolate homologues, for immunological screening and for positive
CC selection methods.
XX
SQ Sequence 1053 BP; 323 A; 198 C; 226 G; 306 T; 0 other;

Query Match 42.9%; Score 497.8; DB 21; Length 1053;
Best Local Similarity 68.9%; Pred. No. 6.5e-121;
Matches 730; Conservative 0; Mismatches 317; Indels 12; Gaps 3;



RESULT 7
AAX90136
ID AAX90136 standard; cDNA; 1041 BP.
XX
AC AAX90136;
XX
DT 20-SEP-1999 (first entry)
XX
DE Coptis japonica norcoclaurine 6-O-methyltransferase encoding cDNA.
XX
KW Coptis japonica; norcoclaurine 6-O-methyltransferase; 6-OMT;
KW secondary metabolite; coclaurine; reticuline; ds.
XX
OS Coptis japonica.
XX
PN JP11178577-A.
XX
PD 06-JUL-1999.
XX
PF 24-DEC-1997; 97JP-0355045.
XX
PR 24-DEC-1997; 97JP-0355045.
XX
PA (MITC) MITSUI PETROCHEM IND CO LTD.
XX
DR WPI; 1999-437310/37.
DR P-PSDB; AAY24396.

XX
PT New norcoclaurine 6-O-methyl-transferase - useful for preparation of
PT secondary plant metabolite
XX
PS Claim 5; Page 12; 27pp; Japanese.
XX
CC The present sequence encodes norcoclaurine 6-O-methyltransferase (6-OMT).
CC Also described is a method for the preparation of the secondary
CC metabolite of a plant derived from coclaurine or reticuline.
XX
SQ Sequence 1041 BP; 320 A; 186 C; 245 G; 290 T; 0 other;

Query Match 18.0%; Score 209.2; DB 20; Length 1041;
Best Local Similarity 53.6%; Pred. No. 6e-45;
Matches 558; Conservative 0; Mismatches 463; Indels 21; Gaps 5;

```
Db        841 GTCGACATTGTCTTAAATGTGCAATCAGAACATCCTTA---TACCAAGATGAGACTGACT 897

Qy        944 ATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAG 1003

Db        898 TTGGATTTGGACATGATGCTCAACACTGGAGGAAAAGAGAGGACTGAAGAGGAATGGAAG 957

Qy       1004 AAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTG 1063

Db        958 AAGCTCATCCATGATGCAGGGTACAAAGGGCATAAGATAACACAAATTACTGCTGTACAA 1017

Qy       1064 TCTCTTATTGAGATCTATCCTT 1085

Db       1018 TCTGTGATTGAGGCTTATCCAT 1039


RESULT 8
AAX89255
ID   AAX89255 standard; cDNA; 1050 BP.
XX
AC   AAX89255;
XX
DT   20-SEP-1999  (first entry)
XX
DE   (S)-3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase encoding cDNA.
XX
KW   (S)-3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase; 4'-OMT;
KW   enzyme; secondary metabolite; reticuline; drug; ss.
XX
OS   Coptis japonika.
XX
PN   JP11178579-A.
XX
PD   06-JUL-1999.
XX
PF   24-DEC-1997;   97JP-0355320.
XX
PR   24-DEC-1997;   97JP-0355320.
XX
PA   (MITC ) MITSUI PETROCHEM IND CO LTD.
XX
DR   WPI; 1999-437312/37.
DR   P-PSDB; AAY27183.
XX
PT   New (S)-3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase  - and
PT   gene encoding it
XX
PS   Claim 5; Fig 12; 28pp; Japanese.
XX
CC   This cDNA encodes a polypeptide having the enzymatic activity of (S)-3'
CC   -hydroxy-N-methylcoclaurine 4'-O-methyltransferase (4'-OMT). A vector
CC   containing the DNA can be used to transform a microbe for the recombinant
CC   preparation of the enzyme. The enzyme can be used to prepare a secondary
CC   metabolite of a plant derived from reticuline which is useful as a raw
CC   material for drugs.
XX
SQ   Sequence 1050 BP; 313 A; 191 C; 252 G; 294 T; 0 other;

  Query Match             15.4%;  Score 178.4;  DB 20;  Length 1050;
  Best Local Similarity   51.3%;  Pred. No. 7.7e-37;
  Matches  528;  Conservative    0;  Mismatches  481;  Indels   21;  Gaps    4;

Qy         57 AAGGTCAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAAT 116

Db         38 AAGCTCAAGCTCATGTGTGGAAAATCATCTATGGTTTTGCAGATTCCCTAGTCCTCCGAT 97

Qy        117 GGATGGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTA 176

Db         98 GTGCAGTGGAACTTGGAATCGTCGACAT------CATTGATAACAACAACCAACCCATGG 151

Qy        177 CTTTTTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGA 236

Db        152 CACTTGCCGATCTGGCATCTAAGCTTCCTGTTTCCGATGTGAATTGCGATAATTTGTATC 211

Qy        237 GCCTCATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACA 296

Db        212 GGATATTACGATACTTGGTGAAAATGGAAATACTGAGAGTGGAAAAATCTGATGATGGTC 271

Qy        297 TAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTT 356

Db        272 AGAAGAAGTACGCGCTTGAACCTATTGCAACATTGCTTTCAAGGAATGCGAAGAGGAGTA 331

Qy        357 TAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGA 416

Db        332 TGGTTCCAATGATTCTTGGAATGACTCAAAAAGATTTTATGACTCCTTGGCATTCAATGA 391

Qy        417 AGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCT 476

Db        392 AGGATGGCTTAAGTGACAATGGTC---ATGCTTTTGAGAAGGCCATGGGAATGACTATAT 448

Qy        477 GGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTG 536

Db        449 GGGAGTACTTGGAAGGACACCCTGATCAAAGCCAATTATTCAATGAAGGCATGGCCGGTG 508

Qy        537 ATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGG---TCTTTGAGGGACTGG 593

Db        509 AAACAAGGCTTCTCACTTCTTCACTCATATCTGGAAGTAGAGATATGTTTCAAGGTATTG 568

Qy        594 AATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTT 653

Db        569 ACTCACTTGTTGATGTTGGTGGAGGAAATGGTACTACTGTCAAGGCCATTTCTGACGCAT 628

Qy        654 TTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAA 713

Db        629 TTCCACATATCAAGTGCACCCTCTTTGATCTCCCTCATGTCATTGCCAATTCCTATGACC 688

Qy        714 GCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAG 773

Db        689 TTCCTAATATTGAACGAATTGGTGGCGACATGTTTAAATCCGTGCCCAGTGCCCAAGCTA 748

Qy        774 TTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAA 833

Db        749 TCATACTCAAGCTAATTTTGCACGATTGGAATGACGAAGACTCGATCAAGATTTTAAAGC 808

Qy        834 ATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTG 893

Db        809 AATGCAGAAATGCAGTGCCA---------AAAGATGGAGGAAAAGTGATTATAGTGGATG 859

Qy        894 TGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTAC 953

Db        860 TGGCATTAGATGAGGAGTCAGACCATGAGCTTAGCAGCACACGATTGATCCTTGATATCG 919

Qy        954 ACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCA 1013

Db        920 ATATGTTGGTGAACACTGGTGGTAAAGAGCGGACTAAAGAGGTTTGGGAGAAAATTGTGA 979

Qy       1014 TGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTG 1073

Db        980 AAAGTGCAGGATTTAGTGGTTGCAAAATCAGGCACATAGCGGCTATACAATCAGTCATTG 1039

Qy       1074 AGATCTATCC 1083

Db       1040 AGGTTTTTCC 1049


RESULT 9
ABX31899
ID   ABX31899 standard; cDNA; 256 BP.
XX
AC   ABX31899;
XX
DT   11-FEB-2003  (first entry)
XX
DE   Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #13956.
XX
KW   Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
```

```
KW   cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW   arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW   complex carbohydrate; gene replacement therapy; immunosuppressive;
KW   antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW   antiasthmatic; vasotropic.
XX
OS   Homo sapiens.
XX
PN   US2002110548-A1.
XX
PD   15-AUG-2002.
XX
PF   11-JUN-2001; 2001US-0878574.
XX
PR   22-NOV-1996;   96US-0753233.
PR   03-DEC-1997;   97US-0984246.
PR   09-SEP-1998;   98US-0149674.
PR   14-JUN-1999;   99US-0333177.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Sullivan F,  Kriz R,  Kumar R;
XX
DR   WPI; 2003-066673/06.
XX
PT   New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT   peptide, for manufacturing complex carbohydrates, or as targets for
PT   screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT   rejection  -
XX
PS   Disclosure; SEQ ID NO 13958; 6pp; English.
XX
CC   The invention relates to a composition comprising a human GDP-mannose
CC   4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
CC   GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
CC   in a mammalian subject and for treating or ameliorating diseases affected
CC   by the level of cellular fucosylation or diseases affected by the
CC   fucosylation of glycoconjugates. These diseases include arthritis,
CC   transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC   infection. The GM4,6D peptide or a polynucleotide encoding it is also
CC   useful for manufacturing complex carbohydrates and as targets for
CC   screening small molecule antagonists of the activity of the enzyme. The
CC   polynucleotide is useful in developing an assay for defects in the
CC   enzyme, as well as in gene replacement therapy. Sequences
CC   ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
CC   human GM4,6D peptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 256 BP; 61 A; 49 C; 65 G; 81 T; 0 other;

  Query Match             13.6%;  Score 158.2;  DB 25;  Length 256;
  Best Local Similarity   76.9%;  Pred. No. 9.5e-32;
  Matches  193;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy        515 TTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAAT 574
              ||||||||||| |||||  |||||||  |  ||  |  |||||||  | | | |||||| |
Db          6 TTCAATGAGGCTATGGCAAGTGATTCCCGAATAGTAGACTTGGCACTCAAAAATTGCACT 65

Qy        575 TGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCA 634
              |  ||  |||||||| || || ||||| ||||||||||||||||||||||||| ||| ||
Db         66 TCAGTTTTTGAGGGGCTAGATTCCATGGTGGATGTTGGTGGTGGAACTGGAACCACGGCC 125

Qy        635 AAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTT 694
              |  |||||||||||  || |||||||| |||||||  | ||| | || | |||   ||||
Db        126 AGAATTATCTGTGACGCATTTCCTAAGTTGAAATGTGTTGTGCTTGACCTTCCTCCTGTT 185

Qy        695 GTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGC 754
              || || || ||| |||| | ||| |||||||  |||||||||| |||||||| || |
Db        186 GTAGAGAACTTGACAGGGACCAATAATTTGAGTTTTGTTGGTGGTGACATGTTCAACTCT 245

Qy        755 ATCCCCAAGGC 765
              |||||  | ||
Db        246 ATCCCTCAAGC 256


RESULT 10
ABX23518
ID   ABX23518 standard; cDNA; 266 BP.
XX
AC   ABX23518;
XX
DT   10-FEB-2003  (first entry)
XX
DE   Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #5575.
XX
KW   Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
KW   cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW   arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW   complex carbohydrate; gene replacement therapy; immunosuppressive;
KW   antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW   antiasthmatic; vasotropic.
XX
OS   Homo sapiens.
XX
PN   US2002110548-A1.
XX
PD   15-AUG-2002.
XX
PF   11-JUN-2001; 2001US-0878574.
XX
PR   22-NOV-1996;   96US-0753233.
PR   03-DEC-1997;   97US-0984246.
PR   09-SEP-1998;   98US-0149674.
PR   14-JUN-1999;   99US-0333177.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Sullivan F,  Kriz R,  Kumar R;
XX
DR   WPI; 2003-066673/06.
XX
PT   New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT   peptide, for manufacturing complex carbohydrates, or as targets for
PT   screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT   rejection  -
XX
PS   Disclosure; SEQ ID NO 5577; 6pp; English.
XX
CC   The invention relates to a composition comprising a human GDP-mannose
CC   4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
CC   GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
CC   in a mammalian subject and for treating or ameliorating diseases affected
CC   by the level of cellular fucosylation or diseases affected by the
CC   fucosylation of glycoconjugates. These diseases include arthritis,
CC   transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC   infection. The GM4,6D peptide or a polynucleotide encoding it is also
CC   useful for manufacturing complex carbohydrates and as targets for
CC   screening small molecule antagonists of the activity of the enzyme. The
CC   polynucleotide is useful in developing an assay for defects in the
CC   enzyme, as well as in gene replacement therapy. Sequences
CC   ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
CC   human GM4,6D peptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 266 BP; 91 A; 42 C; 60 G; 73 T; 0 other;

  Query Match             12.3%;  Score 142.8;  DB 25;  Length 266;
  Best Local Similarity   72.1%;  Pred. No. 1.1e-27;
  Matches  186;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy        701 AATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCC 760
              || ||| |||| | ||| |||||||  |||||||||| |||||||| || |  |||||
```

```
Db          2 AACTTGACAGGGACCAATAATTTGAGTTTTGTTGGTGGTGACATGTTCAACTCTATCCCT 61

Qy        761 AAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATG 820

Db         62 CAAGCTGATGCAGTGCTACTAAAGTGGGTTTTACATAATTGGACCGACGAAAATTGCATA 121

Qy        821 AAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTT 880

Db        122 AAGATCCTGCAAAAGTGTAGAGATTCTATTTCAAGCAAAGGCAACAGTGGAAAAGTGATT 181

Qy        881 GTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTC 940

Db        182 ATCATAGATGCCGTAATAAATGAGAAGCTAGATGACCCGGATATGACACAAACAAAGCTT 241

Qy        941 CTTATGGATGTACACATG 958

Db        242 AGTTTGGACATTATTATG 259


RESULT 11
ABX19667
ID   ABX19667 standard; cDNA; 381 BP.
XX
AC   ABX19667;
XX
DT   10-FEB-2003  (first entry)
XX
DE   Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #1724.
XX
KW   Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
KW   cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW   arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW   complex carbohydrate; gene replacement therapy; immunosuppressive;
KW   antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW   antiasthmatic; vasotropic.
XX
OS   Homo sapiens.
XX
PN   US2002110548-A1.
XX
PD   15-AUG-2002.
XX
PF   11-JUN-2001; 2001US-0878574.
XX
PR   22-NOV-1996;   96US-0753233.
PR   03-DEC-1997;   97US-0984246.
PR   09-SEP-1998;   98US-0149674.
PR   14-JUN-1999;   99US-0333177.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Sullivan F,  Kriz R,  Kumar R;
XX
DR   WPI; 2003-066673/06.
XX
PT   New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT   peptide, for manufacturing complex carbohydrates, or as targets for
PT   screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT   rejection  -
XX
PS   Disclosure; SEQ ID NO 1726; 6pp; English.
XX
CC   The invention relates to a composition comprising a human GDP-mannose
CC   4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
CC   GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
CC   in a mammalian subject and for treating or ameliorating diseases affected
CC   by the level of cellular fucosylation or diseases affected by the
CC   fucosylation of glycoconjugates. These diseases include arthritis,
CC   transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC   infection. The GM4,6D peptide or a polynucleotide encoding it is also
CC   useful for manufacturing complex carbohydrates and as targets for
CC   screening small molecule antagonists of the activity of the enzyme. The
CC   polynucleotide is useful in developing an assay for defects in the
CC   enzyme, as well as in gene replacement therapy. Sequences
CC   ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
CC   human GM4,6D peptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 381 BP; 109 A; 88 C; 69 G; 115 T; 0 other;

  Query Match             12.3%;  Score 142.2;  DB 25;  Length 381;
  Best Local Similarity   66.9%;  Pred. No. 1.8e-27;
  Matches  236;  Conservative    0;  Mismatches  108;  Indels    9;  Gaps    2;

Qy         21 CATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAAC 80

Db         25 CTTCAATGAATAACCAAAAAGAAATTGAGCTCTTTGAGGGCCAATCTCTTCTGTACATGC 84

Qy         81 ATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCG 140

Db         85 AGCTATATGGGCACCTAAGACCTATGTGTCTTAAGTGGGCTGTTCAACTAGGTATTCCAG 144

Qy        141 ACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTC 200

Db        145 ACATAAT------ACAGAACCATGCCAAACCCATTTCTCTTTCTGACTTGGTCTCTACTC 198

Qy        201 TACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACA 260

Db        199 TTCAAATTCCACCAGCTAACGCTGCTTTTGTGCAGCGGTTCATGCGCTTCTTGGCACACA 258

Qy        261 ATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGA---AGCATATGCTCTCACTG 317

Db        259 ATGGAATCTTTGAGATCCATGAGAGCCAAGAAGATCATGAACTAACATATGCTCTAACCC 318

Qy        318 CTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTT 370

Db        319 CTGCATCAAAGCTTCTTGTCAATAGTAGTGATCATTGTCTATCTCCAATGGTT 371


RESULT 12
AAX25200
ID   AAX25200 standard; cDNA; 1412 BP.
XX
AC   AAX25200;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Maize caffeic O-methyltransferase cDNA.
XX
KW   Maize; corn; caffeic O-methyltransferase; lignin; transgenic plant;
KW   ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             139..1266
FT                   /*tag=  a
FT                   /transl_except= (pos:817..819, aa:Ala)
XX
PN   WO9910498-A2.
XX
PD   04-MAR-1999.
XX
PF   24-AUG-1998;   98WO-US17519.
XX
PR   12-MAY-1998;   98US-0076851.
PR   27-AUG-1997;   97US-0057082.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bowen BA,  Helentjaris TG,  Wang X;
XX
DR   WPI; 1999-204667/17.
DR   P-PSDB; AAY05661.
```

```
XX
PT   Nucleic acids encoding plant lignin biosynthesis enzymes - used to
PT   transform plants to modulate lignin biosynthesis
XX
PS   Claim 2; Page 119-121; 166pp; English.
XX
CC   This DNA sequence encodes a caffeic O-methyltransferase (see AAY05661)
CC   of maize.  A polynucleotide having this sequence can be amplified
CC   from a cDNA library prepared from premeiotic to uninucleate tassel
CC   from maize A632 using the primer pair given in AAX25225 and
CC   AAX25226.  The invention provides methods and compositions relating
CC   to altering lignin biosynthesis and/or the lignin composition of
CC   plants.  Isolated nucleic acids (see AAX25196-216) that code for
CC   proteins (see AAY05657-77) involved in lignin biosynthesis are
CC   claimed.  Also claimed are recombinant expression cassettes, host
CC   cells (especially maize or sorghum), and transgenic plants and
CC   seeds.  The claimed nucleic acids can be used to transform a plant
CC   to modulate lignin biosynthesis.  A claimed method involves
CC   transforming a plant cell with a recombinant expression cassette
CC   comprising a lignin biosynthesis polynucleotide operably linked to
CC   a promoter, growing the plant cell under plant growing conditions,
CC   and inducing expression of the polynucleotide for a time sufficient
CC   to modulate (preferably increase) lignin biosynthesis in the plant.
CC   The plant lignins can be used as chemical feedstock.  Plant
CC   material of increased lignin content can be used as a fuel source,
CC   and in the pulp and paper industry.  Decreased lignin content
CC   improves the digestibility of fodder crops.
XX
SQ   Sequence 1412 BP; 324 A; 400 C; 375 G; 313 T; 0 other;

  Query Match             10.4%;  Score 120.2;  DB 20;  Length 1412;
  Best Local Similarity   54.0%;  Pred. No. 1.8e-21;
  Matches  292;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    2;

Qy        555 TGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTG 614
              | | | | || || |||      |   |||   ||  | || ||  | || || || ||||
Db        734 TTGTGCTGAGGGAGTGCGGCGAGACGTTTCGTGGGATCGACTCGTTGGTTGACGTCGGTG 793

Qy        615 GTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGG 674
              ||||   |||   | | ||     |  |||   |  || || ||  | || || |||| |
Db        794 GTGGCCATGGTGGCGCCGCCGCCACCATCGCCGCCGCCTTCCCCCACCTCAAGTGCAGCG 853

Qy        675 TGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTG 734
              || | || |  ||  | ||||| |     |       |||  ||   || || | |
Db        854 TGCTTGACCTCCCGCACGTTGTCGCCGGTGCTCCGTCTGATGGCAACGTGCAGTTCGTCG 913

Qy        735 GTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTAC 794
                ||  | ||||||  |   | || ||     ||    || ||| | || |||     | |||
Db        914 CAGGCAATATGTTTGAGAGTATTCCACCTGCAACCGCTGTTTTCCTCAAGAAAACTCTAC 973

Qy        795 ATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAG 854
              || | |||  |||| | || ||  | |||||||  | ||||| ||  |||| || ||
Db        974 ATGACTGGGGTGACGATGAGTGTGTCAAGATATTGAAGAATTGCAAGCAAGCCATATCTC 1033

Qy        855 GTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATG 914
                  |  || || ||| |  || | ||  |||| ||  | | | |||
Db       1034 CACGGGATGCAGGTGGGAAGGTAATAATCTTGGA---TGTGGTAGTTGGATATAAACAGT 1090

Qy        915 AGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATG 974
                |  |     | ||   |  |  |||||  ||                   |||| |
Db       1091 CAAACATAAAGCATCAAGAGACACAAGTTATGTTTGATTTGTATATGATGGCGGTTAACG 1150

Qy        975 GAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCT 1034
              ||  ||| |  | ||  |||| ||||||||  ||||||  |||||| || ||| ||  ||
Db       1151 GAGTTGAGCGTGACGAGCAAGAGTGGAAGAAGATCTTCACTGAAGCTGGATTCAAAGACT 1210

Qy       1035 ACAAAATATCTCCCTTCACAGGATATTTGTC---TCTTATTGAGATCTATCCTTGAATAC 1091
              |||||||     ||| |||   ||  || | ||      | || ||| ||||||||||||| |
Db       1211 ACAAAATTCTACCCGTCATTGGTGATGTATCGGTCATCATCGAGGTCTATCCTTGAATGC 1270

Qy       1092 T 1092
              |
Db       1271 T 1271
```

```
RESULT 13
AAD19546
ID   AAD19546 standard; cDNA; 1098 BP.
XX
AC   AAD19546;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Medicago sativa caffeoyl CoA 3-O-methyltransferase (CCOMT) cDNA.
XX
KW   Alfalfa; caffeoyl CoA 3-O-methyltransferase; CCOMT; lignin; woody plant;
KW   forage legume; transgenic plant; paper industry; ss.
XX
OS   Medicago sativa.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1098
FT                   /*tag=  a
FT                   /product= "Alfalfa caffeoyl CoA 3-O-methyltransferase
FT                   (CCOMT) protein"
XX
PN   WO200173090-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US09398.
XX
PR   24-MAR-2000; 2000US-192086P.
XX
PA   (ROBE-) ROBERTS NOBLE FOUND INC SAMMUEL.
XX
PI   Dixon RA,  Guo D;
XX
DR   WPI; 2001-616527/71.
DR   P-PSDB; AAE12021.
XX
PT   Transforming forage legumes for modifying lignin composition and
PT   increasing in vivo digestibility  comprises transforming plants with
PT   lignin biosynthetic enzyme genes under lignification-associated tissue
PT   specific promoter -
XX
PS   Example 1; Fig 6; 53pp; English.
XX
CC   The invention relates to methods for producing forage legumes or woody
CC   plants having altered lignin composition. Methods for transforming
CC   forage legumes with a DNA construct comprises alfalfa caffeic acid 3-O-
CC   methyltransferase (COMT), caffeoyl CoA 3-O-methyltransferase (CCOMT)
CC   enzyme or its fragment under a lignification-associated tissue specific
CC   promoter, resulting in the down regulation of the corresponding
CC   homologous OMT genes either through antisense inhibition or sense
CC   suppression. The methods are useful for down-regulation of the
CC   corresponding homologous OMT genes, gene silencing, reduced OMT activity
CC   levels, reduced lignin content, and modified lignin composition in
CC   transgenic plants, and increased digestibility of transgenic plant
CC   materials in ruminant animals. The expression of CCOMT transgene produces
CC   an increased syringyl lignin to guaiacyl lignin ratio in the transformed
CC   plant and greatly improved forage in vivo digestibility, and for making
CC   lignins with altered dimer bonding patterns. Transforming forage legumes
CC   with OMT enzymes is useful to produce plants having modified lignin
CC   content and composition for direct comparison of the effects of lignin
CC   content and/or composition of forage digestibility. The methods are also
CC   useful for producing plants that are modified to alter lignins and
CC   improve pulping characteristics for the paper industry. The present cDNA
CC   sequence encodes alfalfa caffeoyl CoA 3-O-methyltransferase (CCOMT)
CC   protein.
XX
SQ   Sequence 1098 BP; 324 A; 210 C; 240 G; 324 T; 0 other;

  Query Match             10.2%;  Score 118;  DB 22;  Length 1098;
  Best Local Similarity   52.7%;  Pred. No. 6.2e-21;
  Matches  309;  Conservative    0;  Mismatches  265;  Indels   12;  Gaps    2;
```

```
Qy        459 CCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCA 518
              | |  ||||    |   ||  || |         | ||| ||    | |||||||   ||||
Db        470 CTTATGGAATGACAGCCTTTGAATACCATGGAACAGATCCAAGGTTTAACAAGGTTTTCA 529

Qy        519 ATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGG 578
              |  |||  ||| ||  | | |||   ||    |   |   ||   ||  | |   |
Db        530 ACAAGGGGATGTCTGATCACTCTACCATCACAATGAAGAAAATTCTTGAGACCTACACAG 589

Qy        579 TCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGA 638
              |||||  ||  ||  ||||  |||| ||||| |||||||| ||||||      |   || |
Db        590 GTTTTGAAGGCCTTAAATCTCTTGTTGATGTAGGTGGTGGTACTGGAGCTGTAATTAACA 649

Qy        639 TTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGG 698
              ||          | |||  |  ||  |  ||    ||  ||    ||  ||||  | |
Db        650 CGATTGTCTCAAAATATCCCACTATAAAGGGTATAAATTTTGATTTACCCCATGTCATTG 709

Qy        699 AAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCC 758
              ||  ||     ||          |      ||||||||| |||||||||     | || |
Db        710 AAGATGCTCCATCTTATCCAGGAGTTGAGCATGTTGGTGGAGACATGTTTGTCAGTATTC 769

Qy        759 CCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCA 818
              | |||||||||| ||| |  | |||| | |||  ||| | |||| ||| |  | | |||
Db        770 CAAAGGCTGATGCTGTTTTTATGAAGTGGATTTGTCATGACTGGAGTGATGAGCACTGCT 829

Qy        819 TGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAG 878
              ||||  | ||  | ||  ||  | ||  ||  |  |||          | |  ||||||||
Db        830 TGAAATTTTTGAAGAACTGCTATGAGGCACTGCCAG---------ACAATGGAAAAGTGA 880

Qy        879 TTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGC 938
              ||||   |||     |  |   |    ||||   |||  |  ||  ||      |
Db        881 TTGTGGCAGAATGCATACTTCCAGTGGCTCCAGATTCAAGCCTGGCCACAAAAGGTGTGG 940

Qy        939 TCCTTATGGATGTACACA---TGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAG 995
              | |  || |||||   ||   ||||  || |  |  ||| |||||||||| | || |||
Db        941 TTCACATTGATGTGATCATGTTGGCTCATAATCCTGGTGGGAAAGAGAGAACACAAAAAG 1000

Qy        996 ATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAAT 1041
              | |   || | ||   ||  |  || || ||||||  | |  |||| |
Db       1001 AGTTTGAGGATCTTGCCAAAGGTGCTGGATTCCAAGGTTTCAAAGT 1046
```

RESULT 14
AAQ77692
ID AAQ77692 standard; cDNA to mRNA; 1341 BP.
XX
AC AAQ77692;
XX
DT 25-MAR-2003 (updated)
DT 10-MAY-1995 (first entry)
XX
DE Alfalfa COMT clone pCOMT1.
XX
KW Alfalfa; tobacco; Nicotiana tabacum; crop improvement; COMT;
KW caffeic-acid-3-O-methyltransferase; lignin; pCOMT1; antisense; ds.
XX
OS Medicago sativa.
XX

FH	Key	Location/Qualifiers
FT	CDS	55..1152
FT		/*tag= a
FT	misc_feature	10..541
FT		/*tag= b
FT		/note= "the inverted complement of this portion
FT		of the COMT gene was the antisense fragment
FT		used to reduce tobacco lignin content"
FT	misc_feature	541..1016
FT		/*tag= c
FT		/note= "the inverted complement of this portion
FT		of the COMT gene was the antisense fragment
FT		used to reduce alfalfa lignin content"

XX
PN WO9423044-A1.
XX
PD 13-OCT-1994.
XX
PF 31-MAR-1994; 94WO-US03356.
XX
PR 02-APR-1993; 93US-0045263.
XX
PA (ROBE-) ROBERTS NOBLE FOUND INC SAMMUEL.
XX
PI Dixon RA, Ni W;
XX
DR WPI; 1994-333204/41.
DR P-PSDB; AAR63203.
XX
PT Reducing lignin content in plants - by transfection with an
PT anti-sense gene for caffeic acid 3-O-methyl-transferase gene.
XX
PS Disclosure; Page 21-23; 39pp; English.
XX
CC Alfalfa COMT clone pCOMT1, obtained from a cDNA library in lambda
CC ZAP II, was the basis of antisense genes used to reduce the lignin
CC content of e.g. tobacco and alfalfa.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1341 BP; 402 A; 252 C; 280 G; 407 T; 0 other;

Query Match 10.2%; Score 118; DB 15; Length 1341;
Best Local Similarity 52.7%; Pred. No. 6.6e-21;
Matches 309; Conservative 0; Mismatches 265; Indels 12; Gaps 2;

```
Qy        459 CCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCA 518
              | |  ||||    |   ||  || |         | ||| ||    | |||||||   ||||
Db        524 CTTATGGAATGACAGCCTTTGAATACCATGGAACAGATCCAAGGTTTAACAAGGTTTTCA 583

Qy        519 ATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGG 578
              |  |||  ||| ||  | | |||   ||    |   |   ||   ||  | |   |
Db        584 ACAAGGGGATGTCTGATCACTCTACCATCACAATGAAGAAAATTCTTGAGACCTACACAG 643

Qy        579 TCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGA 638
              |||||  ||  ||  ||||  |||| ||||| |||||||| ||||||      |   || |
Db        644 GTTTTGAAGGCCTTAAATCTCTTGTTGATGTAGGTGGTGGTACTGGAGCTGTAATTAACA 703

Qy        639 TTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGG 698
              ||          | |||  |  ||  |  ||    ||  ||    ||  ||||  | |
Db        704 CGATTGTCTCAAAATATCCCACTATAAAGGGTATAAATTTTGATTTACCCCATGTCATTG 763

Qy        699 AAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCC 758
              ||  ||     ||          |      ||||||||| |||||||||     | || |
Db        764 AAGATGCTCCATCTTATCCAGGAGTTGAGCATGTTGGTGGAGACATGTTTGTCAGTATTC 823

Qy        759 CCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCA 818
              | |||||||||| ||| |  | |||| | |||  ||| | |||| ||| |  | | |||
Db        824 CAAAGGCTGATGCTGTTTTTATGAAGTGGATTTGTCATGACTGGAGTGATGAGCACTGCT 883

Qy        819 TGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAG 878
              ||||  | ||  | ||  ||  | ||  ||  |  |||          | |  ||||||||
Db        884 TGAAATTTTTGAAGAACTGCTATGAGGCACTGCCAG---------ACAATGGAAAAGTGA 934

Qy        879 TTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGC 938
              ||||   |||     |  |   |    ||||   |||  |  ||  ||      |
Db        935 TTGTGGCAGAATGCATACTTCCAGTGGCTCCAGATTCAAGCCTGGCCACAAAAGGTGTGG 994

Qy        939 TCCTTATGGATGTACACA---TGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAG 995
              | |  || |||||   ||   ||||  || |  |  ||| |||||||||| | || |||
Db        995 TTCACATTGATGTGATCATGTTGGCTCATAATCCTGGTGGGAAAGAGAGAACACAAAAAG 1054

Qy        996 ATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAAT 1041
              | |   || | ||   ||  |  || || ||||||  | |  |||| |
Db       1055 AGTTTGAGGATCTTGCCAAAGGTGCTGGATTCCAAGGTTTCAAAGT 1100
```

RESULT 15

```
AAC40707
ID   AAC40707 standard; DNA; 1267 BP.
XX
AC   AAC40707;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 29270.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
```

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.8%; Score 114; DB 21; Length 1267;
Best Local Similarity 54.9%; Pred. No. 7.3e-20;
Matches 225; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

```
Qy        494 GACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAAC 553
              ||||||  ||  |||||||   || ||  | | |||| ||    ||||    ||    |
Db        578 GACCCTAGATTCAACAAGGTCTTTAACAATGGAATGTCTAACCATTCCACAATCACCATG 637

Qy        554 TTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGT 613
                |  | ||   ||   | ||  ||  ||||| ||| ||   ||  | || |||||||||
Db        638 AAGAAGATTCTTGAGACCTATAAGGGTTTTGAAGGATTGACTTCTTTGGTTGATGTTGGT 697

Qy        614 GGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATG 673
              |||||  | |||    |||   || || ||         |  ||||| || ||| ||||
Db        698 GGTGGCATTGGTGCTACACTCAAAATGATTGTCTCCAAGTACCCTAATCTTAAAGGCATC 757

Qy        674 GTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTT 733
```

```
                ||  ||  |  |||  ||||  |  |||  ||    |             |  |      ||||
Db        758 AACTTTGATCTCCCACATGTCATTGAAGATGCTCCTTCTCATCCTGGTATTGAGCATGTT 817

Qy        734 GGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTA 793
              ||  ||  ||  ||||||  |  |  ||||  ||  |  ||||||  |  |  |  ||||  |  |  |
Db        818 GGAGGAGATATGTTTGTAAGTGTCCCTAAAGGTGATGCCATATTCATGAAGTGGATATGT 877

Qy        794 CATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCA 853
              |||  |  ||||  ||||  |  |||||  ||||  |  ||  ||||  ||  |  ||  |  ||  ||
Db        878 CATGACTGGAGTGACGAACATTGCGTGAAATTCTTGAAAAACTGCTACGAGTCACTTCCA 937

Qy        854 GGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGA 903
              |  ||  |  |||      |  ||  ||     |||  |  |  |  ||  |   ||
Db        938 GAGGATGGAAAAGTGATATTAGCAGAGTGTATACTTCCAGAGACACCAGA 987
```

Search completed: November 9, 2003, 23:45:31
Job time : 432 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 9, 2003, 23:36:54 ; Search time 102 Seconds
(without alignments)
5019.651 Million cell updates/sec

Title: US-09-868-547-3
Perfect score: 1160
Sequence: 1 gtttgctattatggcttctt.....ataaaagttatctttgtgtt 1160

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	114	9.8	1314	4	US-09-500-569-17	Sequence 17, Appl
	2	109.8	9.5	1368	2	US-08-204-288-1	Sequence 1, Appli
	3	109.8	9.5	1503	2	US-08-715-325-1	Sequence 1, Appli
	4	109	9.4	1507	4	US-09-453-323-1	Sequence 1, Appli
	5	108	9.3	1244	2	US-08-204-288-3	Sequence 3, Appli
	6	107.8	9.3	1370	4	US-09-500-569-3	Sequence 3, Appli
	7	97	8.4	1430	2	US-08-204-288-4	Sequence 4, Appli
	8	95.4	8.2	1195	4	US-09-500-569-11	Sequence 11, Appl
	9	94.6	8.2	1380	3	US-08-991-677-5	Sequence 5, Appli
	10	92.6	8.0	1630	4	US-09-615-192A-107	Sequence 107, App
	11	89.6	7.7	381	4	US-09-615-192A-174	Sequence 174, App
	12	86.2	7.4	661	4	US-09-615-192A-127	Sequence 127, App
	13	86.2	7.4	744	2	US-08-975-316-51	Sequence 51, Appl
	14	86.2	7.4	744	4	US-09-615-192A-51	Sequence 51, Appl
	15	82.6	7.1	1471	2	US-08-204-288-6	Sequence 6, Appli
	16	81.6	7.0	1525	1	US-08-186-833-3	Sequence 3, Appli
	17	74.6	6.4	1308	4	US-09-500-569-13	Sequence 13, Appl
	18	70.2	6.1	1458	4	US-09-500-569-15	Sequence 15, Appl
	19	69.4	6.0	1438	2	US-08-845-742-1	Sequence 1, Appli
	20	61.8	5.3	1342	4	US-09-500-569-9	Sequence 9, Appli
	21	60.6	5.2	470	4	US-09-615-192A-189	Sequence 189, App
	22	59	5.1	438	4	US-09-615-192A-187	Sequence 187, App
	23	57.8	5.0	1036	4	US-09-615-192A-191	Sequence 191, App
c	24	57.2	4.9	7218	1	US-08-232-463-14	Sequence 14, Appl
	25	54.2	4.7	682	4	US-09-615-192A-192	Sequence 192, App
	26	53.6	4.6	760	4	US-09-500-569-7	Sequence 7, Appli
	27	47.8	4.1	849	4	US-09-500-569-25	Sequence 25, Appl
	28	45.8	3.9	1314	4	US-09-500-569-5	Sequence 5, Appli
	29	40.6	3.5	1956	4	US-08-559-896B-1	Sequence 1, Appli
	30	40.6	3.5	12886	4	US-09-453-702B-14	Sequence 14, Appl
	31	40.2	3.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	32	38.2	3.3	289	3	US-09-007-005-17	Sequence 17, Appl
	33	38.2	3.3	289	3	US-09-244-796-17	Sequence 17, Appl
c	34	38	3.3	202001	4	US-09-734-674-3	Sequence 3, Appli
c	35	37.6	3.2	319608	4	US-09-539-333D-1	Sequence 1, Appli
c	36	37.6	3.2	319608	4	US-09-679-409-1	Sequence 1, Appli
	37	37.2	3.2	549	4	US-09-107-532A-2800	Sequence 2800, Ap
	38	36.8	3.2	4892	4	US-09-620-312D-1065	Sequence 1065, Ap
	39	36.6	3.2	993	4	US-09-134-001C-974	Sequence 974, App
	40	36.6	3.2	5361	3	US-08-973-462-2	Sequence 2, Appli
	41	36.6	3.2	6152	3	US-08-973-462-1	Sequence 1, Appli
	42	35.2	3.0	4529	2	US-08-449-645A-16	Sequence 16, Appl
	43	35.2	3.0	4529	2	US-08-702-367A-16	Sequence 16, Appl
	44	35.2	3.0	4529	5	PCT-US95-04681-16	Sequence 16, Appl
	45	35	3.0	35100	1	US-08-306-691B-19	Sequence 19, Appl

ALIGNMENTS

```
RESULT 1
US-09-500-569-17
; Sequence 17, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Shen, Jennie
;  TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
;  FILE REFERENCE: BB1327 US NA
;  CURRENT APPLICATION NUMBER: US/09/500,569
;  CURRENT FILING DATE: 2000-02-09
;  EARLIER APPLICATION NUMBER: 60/119,587
;  EARLIER FILING DATE: 1999-February-10
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
;   LENGTH: 1314
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (472)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1156)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1180)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1262)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1302)
US-09-500-569-17

  Query Match             9.8%;  Score 114;  DB 4;  Length 1314;
  Best Local Similarity   54.5%;  Pred. No. 8e-22;
  Matches  228;  Conservative    0;  Mismatches  190;  Indels    0;  Gaps    0;

Qy        448 ATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAA 507
              |||| |    | |  ||||   |  ||  || |     |  || ||    | |||
Db        530 ATTTAACAAAGCATATGGAATGACAGCCTTTGAATACCATGGAACGGATCCAAGGTTTAA 589

Qy        508 CAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGA 567
              ||||   |||||  |||  ||||||  | | |||   ||    |    |   ||   ||
Db        590 CAAGGTTTTCAACAAGGGGATGGCTGATCACTCTACCATTACAATGAAGAAAATTCTTGA 649

Qy        568 TTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAAT 627
```

```
Db        650 GACCTACACAGGCTTTGAGGGACTTAAATCCCTGGTTGATGTTGGTGGAGGAACTGGAGC 709

Qy        628 CACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCC 687

Db        710 TGTAGTCAACATGATTGTCTCAAAGTATCCCACTATTAAGGGCATTAATTTTGATTTGCC 769

Qy        688 AAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTT 747

Db        770 CCATGTCATTGAAGATGCCCCATCTTATCCAGGAGTGGAACATGTTGGTGGAGATATGTT 829

Qy        748 TAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGA 807

Db        830 TGTCAGTGTTCCAAAAGCTGATGCTATTTTTATGAAGTGGATTTGCCACGATTGGAGTGA 889

Qy        808 CAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAA 865

Db        890 TGAGCACTGCTTGAAGTTTTTGAAGAACTGCTATGAGGCACTACCAGATAATGGGAAA 947


RESULT 2
US-08-204-288-1
; Sequence 1, Application US/08204288
; Patent No. 5959178
;  GENERAL INFORMATION:
;    APPLICANT:  VAN DOORSSELAERE, Jan
;    APPLICANT:  FRITIG, Bernard J.M.
;    APPLICANT:  INZE, Dirk G.
;    APPLICANT:  JOUANIN, Lise
;    APPLICANT:  KNIGHT, Mary E.
;    APPLICANT:  VAN MONTAGU, Marc
;    APPLICANT:  LEGRAND, Michel
;    TITLE OF INVENTION:  MODIFICATION OF LIGNIN SYNTHESIS IN
;    TITLE OF INVENTION:  PLANTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,288
;      FILING DATE:  10-MAR-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9119279.9
;      FILING DATE:  10-SEP-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/01460
;      FILING DATE:  09-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOKULIS, Paul N.
;      REGISTRATION NUMBER:  16,773
;      REFERENCE/DOCKET NUMBER:  206860/SEE36543/UST
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1368 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-204-288-1
```

```
  Query Match             9.5%;  Score 109.8;  DB 2;  Length 1368;
  Best Local Similarity   49.6%;  Pred. No. 1.1e-20;
  Matches  313;  Conservative    0;  Mismatches  312;  Indels    6;  Gaps    1;

Qy        448 ATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAA 507

Db        507 ATTTAACAAGGCCTATGGGATGACTGCATTTGAATATCATGGCACGGATCCAAGATTCAA 566

Qy        508 CAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGA 567

Db        567 CAAGGTCTTCAATAAGGGAATGTCTGACCACTCTACCATTACCATGAAGAAGATTCTTGA 626

Qy        568 TTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAAT 627

Db        627 GACCTACAAAGGCTTTGAAGGCCTCACATCCTTGGTGGATGTTGGTGGTGGGACTGGAGC 686

Qy        628 CACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCC 687

Db        687 TGTCGTTAACACCATCGTCTCTAAATACCCTTCAATTAAGGGCATTAACTTTGATTTGCC 746

Qy        688 AAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTT 747

Db        747 CCACGTCATTGAGGATGCCCCATCTTATCCCGGTGTGGAGCATGTTGGTGGGGACATGTT 806

Qy        748 TAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGA 807

Db        807 TGTTAGTGTGCCCAAAGCAGATGCCGTTTTCATGAAGTGGATATGCCATGATTGGAGCGA 866

Qy        808 CAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAAC 867

Db        867 CGCACACTGCTTAAAATTCTTGAAGAATTGCTATGACGCCTTGCCGGAAAACGGCAAGGT 926

Qy        868 AGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTAC 927

Db        927 GATACTTGTTGAGTGCATTCTTCCCGTGGCTCCTGACACAAGCCTTGCCACCAAGGG--- 983

Qy        928 TGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAA 987

Db        984 ---AGTCGTTCACATTGATGTTATCATGCTGGCGCACAACCCCGGTGGGAAAGAGAGGAC 1040

Qy        988 AGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCC 1047

Db       1041 CGAAAAGGAATTTGAGGGCTTAGCTAAGGGAGCTGGCTTTCAAGGTTTTGAAGTGATGTG 1100

Qy       1048 CTTCACAGGATATTTGTCTCTTATTGAGATC 1078

Db       1101 CTGTGCATTCAACACACATGTCATTGAACTC 1131


RESULT 3
US-08-715-325-1
; Sequence 1, Application US/08715325
; Patent No. 5886243
;  GENERAL INFORMATION:
;    APPLICANT:  Chiang, Vincent L
;    APPLICANT:  Tsai, Chung-Jui
;    APPLICANT:  Podila, Gopi
;    TITLE OF INVENTION:  Genetic Engineering of Wood Color
;    TITLE OF INVENTION:  in Plants
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael, Best & Friedrich
;      STREET:  100 E. Wisconsin Avenue
;      CITY:  Milwaukee
;      STATE:  Wisconsin
;      COUNTRY:  United States of America
;      ZIP:  53202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS DOS 6.22
;      SOFTWARE:  WordPerfect for Windows 5.2
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/715,325
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US60/007727
;      FILING DATE:  30 NOV 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gemignani, Joseph A
;      REGISTRATION NUMBER:  19,482
;      REFERENCE/DOCKET NUMBER:  66040/9627
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (414) 271-6560
;      TELEFAX:  (414) 277-0656
;      TELEX:  262057
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1503 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    ORIGINAL SOURCE:
;      ORGANISM:  Populus Tremuloides
;      DEVELOPMENTAL STAGE:  four year old sapling
;      DEVELOPMENTAL STAGE:  undergoing lignification in summer
;      TISSUE TYPE:  secondary xylem
;    IMMEDIATE SOURCE:
;      LIBRARY:  cDNA to total mRNA
;      CLONE:  PtOMT1
US-08-715-325-1

  Query Match             9.5%;  Score 109.8;  DB 2;  Length 1503;
  Best Local Similarity   49.6%;  Pred. No. 1.2e-20;
  Matches  313;  Conservative    0;  Mismatches  312;  Indels    6;  Gaps    1;

Qy        448 ATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAA 507
              |||| |     |||  || |   ||   | || | |     |  || ||    ||  ||
Db        523 ATTTAACAAGGCCTATGGGATGACTGCATTTGAATATCATGGCACGGATCCAAGATTCAA 582

Qy        508 CAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGA 567
              ||||   |||||  ||| |||| ||    | |||   ||     |   | | ||   ||
Db        583 CAAGGTCTTCAACAAGGGAATGTCTGACCACTCTACCATTACCATGAAGAAGATTCTTGA 642

Qy        568 TTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAAT 627
                | |    | ||||||  || ||     ||| | ||||||||||||||||||  ||||||
Db        643 GACCTACAAAGGCTTTGAAGGCCTCACGTCCTTGGTGGATGTTGGTGGTGGGACTGGACC 702

Qy        628 CACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCC 687
              |    | || |  |||          |  |||     | ||  ||||     || ||  |  ||
Db        703 CGTCGTTAACACCATCGTCTCTAAATACCCTTCAATCAAGGGCATTAACTTCGATCTGCC 762

Qy        688 AAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTT 747
                | ||  | ||  ||    ||        |    ||      |||||||| ||||||||
Db        763 CCACGTCATTGAGGATGCCCCATCTTATCCCGGAGTGGAGCATGTTGGTGGCGACATGTT 822

Qy        748 TAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGA 807
              |    |  | ||||| || ||||| ||| |  | |||| | | |  ||| ||||||  ||
Db        823 TGTTAGTGTGCCCAAAGCAGATGCCGTTTTCATGAAGTGGATATGCCATGATTGGAGCGA 882

Qy        808 CAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAAC 867
              |  | | ||| | ||  | ||  | |||||  | || ||   |  | |   |  ||||
Db        883 CGCCCACTGCTTAAAATTCTTGAAGAATTGCTATGACGCGTTGCCGGAAAACGGCAAGGT 942

Qy        868 AGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTAC 927
              |   || |    |||   | | |||      || |  |    ||    | | |
Db        943 GATACTTGTTGAGTGCATTCTTCCCGTGGCTCCTGACACAAGCCTTGCCACCAAGGG--- 999

Qy        928 TGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAA 987
              | |   || | ||   |   |    ||||    |      ||| |||||||| |
Db       1000 ---AGTCGTGCACGTTGATGTCATCATGCTGGCGCACAACCCCGGTGGGAAAGAGAGGAC 1056

Qy        988 AGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCC 1047
              ||  | || |   ||     |    | || ||| || |||||| | |   || ||
Db       1057 CGAGAAGGAATTTGAGGGCTTAGCTAAGGGAGCTGGCTTCCAAGGTTTTGAAGTAATGTG 1116

Qy       1048 CTTCACAGGATATTTGTCTCTTATTGAGATC 1078
              ||   ||    |      | | |||||  ||
Db       1117 CTGTGCATTCAACACACATGTCATTGAATTC 1147


RESULT 4
US-09-453-323-1
; Sequence 1, Application US/09453323
; Patent No. 6441272
; GENERAL INFORMATION:
;  APPLICANT:  Ye, Zheng-Hua
;  TITLE OF INVENTION:  MODIFICATION OF LIGNIN CONTENT AND COMPOSITION IN
;  TITLE OF INVENTION:  PLANTS
;  FILE REFERENCE:  235.00090101
;  CURRENT APPLICATION NUMBER: US/09/453,323
;  CURRENT FILING DATE:  1999-12-02
;  PRIOR APPLICATION NUMBER:  60/110,676
;  PRIOR FILING DATE:  1998-12-02
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1507
;   TYPE: DNA
;   ORGANISM: Nicotiana tabacum
US-09-453-323-1

  Query Match             9.4%;  Score 109;  DB 4;  Length 1507;
  Best Local Similarity   53.1%;  Pred. No. 2e-20;
  Matches  232;  Conservative    0;  Mismatches  205;  Indels    0;  Gaps    0;

Qy        459 CCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCA 518
              |||  ||||   |   |  || |     | ||| ||   ||  |||||    ||||
Db        475 CCTATGGAATGACAGCATTTGAGTACCATGGCACAGATCCAAGATTCAACAAAGTGTTCA 534

Qy        519 ATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGG 578
              |    | |||| ||  | | ||    ||||  |    |  | ||   ||   | |    |
Db        535 ACCGTGGAATGTCTGATCACTCCACTATGTCAATGAAGAAGATTCTTGAGGACTACAAAG 594

Qy        579 TCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGA 638
              |||||  || ||  | |||||||| ||||||||||||||||||  ||    ||||  || |
Db        595 GATTTGAAGGCCTAAATTCCATTGTTGATGTTGGTGGTGGAACGGGTGCTACAGTTAACA 654

Qy        639 TTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGG 698
              | ||          | |||     | ||  ||||    || ||     ||| ||||  | |
Db        655 TGATTGTCTCTAAATATCCCTCTATTAAGGGCATTAACTTTGATTTGCCACATGTAATTG 714

Qy        699 AAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCC 758
              | ||    ||       |  | | |        |||||||| |||||||||     |  | |
Db        715 GAGATGCTCCAACTTACCCCGGTGTCGAGCACGTTGGTGGCGACATGTTTGCTAGTGTGC 774

Qy        759 CCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCA 818
              | ||  ||  |||||  ||  |  | |||| | |||  ||| ||||||  ||  |  |||||
Db        775 CAAAAGCAGATGCCATTTTCATGAAGTGGATTTGTCATGATTGGAGCGATGAGCATTGCC 834

Qy        819 TGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAG 878
              | ||  | ||  | ||||||  | |||||  |  | |   |  | ||     ||  | ||
Db        835 TAAAATTCTTGAAGAATTGCTATGAAGCACTACCTGCAAATGGGAAGGTGATAATTGCAG 894

Qy        879 TTGTCATAGATACTGTG 895
                ||||  |  | | |
Db        895 AGTGCATACTTCCAGAG 911


RESULT 5
US-08-204-288-3
; Sequence 3, Application US/08204288
; Patent No. 5959178
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  VAN DOORSSELAERE, Jan
;    APPLICANT:  FRITIG, Bernard J.M.
;    APPLICANT:  INZE, Dirk G.
;    APPLICANT:  JOUANIN, Lise
;    APPLICANT:  KNIGHT, Mary E.
;    APPLICANT:  VAN MONTAGU, Marc
;    APPLICANT:  LEGRAND, Michel
;    TITLE OF INVENTION:  MODIFICATION OF LIGNIN SYNTHESIS IN
;    TITLE OF INVENTION:  PLANTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,288
;      FILING DATE:  10-MAR-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9119279.9
;      FILING DATE:  10-SEP-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/01460
;      FILING DATE:  09-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOKULIS, Paul N.
;      REGISTRATION NUMBER:  16,773
;      REFERENCE/DOCKET NUMBER:  206860/SEE36543/UST
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1244 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-204-288-3

  Query Match             9.3%;  Score 108;  DB 2;  Length 1244;
  Best Local Similarity   52.6%;  Pred. No. 3.5e-20;
  Matches  259;  Conservative    0;  Mismatches  230;  Indels    3;  Gaps    1;

Qy        404 TGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTA 463

Db        268 TGGTACCACTTAAAAGATGCAGTACTAGATGGTGGAATCCCATTCAACAAGGCCTATGGA 327

Qy        464 GGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAG 523

Db        328 ATGACAGCATTTGAGTACCATGGC---ACAGATCCAAGATTCAACAAAGTTTTCAACCGT 384

Qy        524 GCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTT 583

Db        385 GGAATGTCTGATCACTCCACTATGTCAATGAAAAAGATTCTTGAGGACTACAAAGGATTT 444

Qy        584 GAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATC 643

Db        445 GAAGGCCTAAATTCCATTGTCGATGTTGGTGGTGGAACTGGCGCTACTGTTAACATGATT 504

Qy        644 TGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAAT 703

Db        505 GTCTCCAAACATCCCTCTATTAAGGGTATTAACTTTGATTTACCACATGTTATTGGAGAT 564

Qy        704 TTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAG 763

Db        565 GCTCCAGCTTACCCTGGTGTCGAGCACGTTGGTGGCGACATGTTTGCCAGTGTGCCAAAA 624

Qy        764 GCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAG 823

Db        625 GCAGATGCCATTTTCATGAAGTGGATTTGTCATGATTGGAGCGACGAGCATTGCCTAAAA 684

Qy        824 ATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTC 883

Db        685 TTCTTGAAGAATTGCTATGAAGCACTACCTGCAAATGGGAAGGTGATAATAGCGGAGTGC 744

Qy        884 ATAGATACTGTG 895

Db        745 ATACTTCCAGAG 756

RESULT 6
US-09-500-569-3
; Sequence 3, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Shen, Jennie
;  TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
;  FILE REFERENCE: BB1327 US NA
;  CURRENT APPLICATION NUMBER: US/09/500,569
;  CURRENT FILING DATE: 2000-02-09
;  EARLIER APPLICATION NUMBER: 60/119,587
;  EARLIER FILING DATE: 1999-February-10
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
;   LENGTH: 1370
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-500-569-3

  Query Match             9.3%;  Score 107.8;  DB 4;  Length 1370;
  Best Local Similarity   50.9%;  Pred. No. 4.1e-20;
  Matches  315;  Conservative    0;  Mismatches  292;  Indels   12;  Gaps    2;

Qy        434 GAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAA 493

Db        523 GAAGGAGGCATTCCATTTAACAAGGCTTATGGAATGACAGCCTTTGAGTACCATGGAACA 582

Qy        494 GACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAAC 553

Db        583 GATCCAAGGTTTAACAAGGTTTTCAACAAGGGAATGGCTGATCACTCTACCATCACAATG 642

Qy        554 TTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGT 613

Db        643 AAGAAAATTCTTGAGACCTACACAGGTTTTGAGAGTCTTAAATCTCTGGTTGATGTTGGT 702

Qy        614 GGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATG 673

Db        703 GGTGGGACTGGAGCTGTAATCAACATGATTGTCTCAAAGCATCCCACTATTAAGGGCATT 762

Qy        674 GTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTT 733

Db        763 AATTTTGATTTGCCTCATGTCATTGAAGATGCCCCATCTTATCCTGGAGTGGAGCATGTA 822

Qy        734 GGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTA 793

Db        823 GGTGGAGATATGTTTGCGAGTGTTCCGAAAGCTGATGCTATTTTTATGAAGTGGATTTGC 882

Qy        794 CATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCA 853

Db        883 CACGATTGGAGTGATGAGCACTGCTTGAAGTTTTTGAAGAACTGCTACGAGGCACTACCA 942

Qy        854 GGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGAT 913
```

```
Db        943 G---------ACAATGGGAAGGTGATTGTGGCAGAATGCATTCTTCCAGTGGCTCCAGAC 993

Qy        914 GAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACA---TGGCATGTATTATT 970

Db        994 TCTAGCTTGGCCACAAAAGGTGTGGTTCACATCGATGTGATCATGTTGGCACATAATCCA 1053

Qy        971 AATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAA 1030

Db       1054 CGTGGGAAAGAGAGAACAGAGAAAGAGTTTGAGGCTCTGGCCAAAGGGTCTGGATTCCAA 1113

Qy       1031 AGCTACAAAATATCTCCCT 1049

Db       1114 GGTTTCCGAGTTGTTTGCT 1132
```

RESULT 7
US-08-204-288-4
; Sequence 4, Application US/08204288
; Patent No. 5959178
; GENERAL INFORMATION:
; APPLICANT: VAN DOORSSELAERE, Jan
; APPLICANT: FRITIG, Bernard J.M.
; APPLICANT: INZE, Dirk G.
; APPLICANT: JOUANIN, Lise
; APPLICANT: KNIGHT, Mary E.
; APPLICANT: VAN MONTAGU, Marc
; APPLICANT: LEGRAND, Michel
; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,288
; FILING DATE: 10-MAR-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9119279.9
; FILING DATE: 10-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01460
; FILING DATE: 09-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 206860/SEE36543/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1430 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-204-288-4

Query Match 8.4%; Score 97; DB 2; Length 1430;
Best Local Similarity 52.9%; Pred. No. 3.8e-17;
Matches 231; Conservative 0; Mismatches 205; Indels 1; Gaps 1;

```
Qy        459 CCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCA 518

Db        528 CCTATGGAATGCAAGCATTTGAGTACCATGGCACAGATCCAAGATTCAACAAAGTGTTCA 587

Qy        519 ATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGG 578

Db        588 ACCGTGGAATGTCTGATCACTCCACTATGTCAATGAAGAAGATTCTTGAGGACTACAAAG 647

Qy        579 TCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGA 638

Db        648 GATTTGAAGGCCTAAATTCCATTGTTGATGTTGGTGGTGGAACGGGTGCTACTGTTAACA 707

Qy        639 TTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGG 698

Db        708 TGATTGTCTCTAAATATCCCTCTATTAAGGGCATTAACTTTGATTTGCCACATGTAATTG 767

Qy        699 AAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCC 758

Db        768 GAGATGCTCCAACTTACCCCGGTGTCGAGCACGTTGGTGGCGACATGTTTGCTAGTGT-G 826

Qy        759 CCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCA 818

Db        827 CCAAAGCAGATGCCATTTTCATGAAGTGGATTTGTCATGATTGGAGCGATGAGCATTGCC 886

Qy        819 TGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAG 878

Db        887 TAAAATTCTTGAAGAATTGCTATGAAGCACTACCTGCAAATGGGAAGGTGATAATTGCAG 946

Qy        879 TTGTCATAGATACTGTG 895

Db        947 AGTGCATACTTCCAGAG 963
```

RESULT 8
US-09-500-569-11
; Sequence 11, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
; FILE REFERENCE: BB1327 US NA
; CURRENT APPLICATION NUMBER: US/09/500,569
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,587
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1195
; TYPE: DNA
; ORGANISM: Glycine max
US-09-500-569-11

Query Match 8.2%; Score 95.4; DB 4; Length 1195;
Best Local Similarity 52.4%; Pred. No. 9.7e-17;
Matches 210; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

```
Qy        464 GGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAG 523

Db        502 GGAAAACACGTGTTTGAATATTCCGACATGAACTCGAGCTTCAATCAGCTTTTCATGGCA 561

Qy        524 GCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTT 583

Db        562 GCTATGACAAACCGTGCAACTTTAATAATGAAGAAGATTGTTGAATCCTACAAGGGGTTT 621

Qy        584 GAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATC 643

Db        622 GAGCACCTCAATAGCCTGGTGGACGTTGGAGGTGGCCTTGGTGTCACACTTAACATAGTC 681

Qy        644 TGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAAT 703
```

```
Db        682 ACTTCTAAATACCCTCACATTAAGGGTATCAATTTTGACTTGCCACATGTCATAGAACAT 741

Qy        704 TTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAG 763

Db        742 GCCTCTACCTATCCTGGTGTTGAGCATGTGGGAGGAGATATGTTTGAAAGTGTGCCACAA 801

Qy        764 GCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAG 823

Db        802 GGAGATGCCATTTTGATGATGTGTGTACTTCATGATTGGAGTGATGAATGGTGCTTGAAG 861

Qy        824 ATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAA 864

Db        862 GTATTAAAGAATTGTTATGCTTCTATTCCTAGTGATGGAAA 902


RESULT 9
US-08-991-677-5
; Sequence 5, Application US/08991677A
; Patent No. 6252135
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent L
;  APPLICANT: Carraway, Daniel T
;  APPLICANT: Smeltzer, Richard H
;  TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
;  FILE REFERENCE: 50617
;  CURRENT APPLICATION NUMBER: US/08/991,677A
;  CURRENT FILING DATE: 1997-12-16
;  EARLIER APPLICATION NUMBER: US 60/033,381
;  EARLIER FILING DATE: 1996-12-16
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 1380
;   TYPE: DNA
;   ORGANISM: Liquidambar styraciflua
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (67)..(1170)
US-08-991-677-5

  Query Match             8.2%;  Score 94.6;  DB 3;  Length 1380;
  Best Local Similarity   50.1%;  Pred. No. 1.7e-16;
  Matches  325;  Conservative    0;  Mismatches  309;  Indels   15;  Gaps    3;

Qy        404 TGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTA 463

Db        490 TGGTACCACTTGACCGAGGCAGTTCTTGAAGGTGGAATTCCATTTAACAAGGCCTATGGA 549

Qy        464 GGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAG 523

Db        550 ATGACAGCATTTGAGTACCATGGC---ACCGATCCCAGATTCAACACAGTTTTCAACAAT 606

Qy        524 GCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTT 583

Db        607 GGAATGTCCAATCATTCGACCATTACCATGAAGAAAATCCTTGAGACTTACAAAGGGTTC 666

Qy        584 GAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATC 643

Db        667 GAGGGACTTGGATCTGTGGTTGATGTTGGTGGTGGCACTGGTGCCCACCTTAACATGATT 726

Qy        644 TGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAAT 703

Db        727 ATCGCTAAATACCCCATGATCAAGGGCATTAACTTCGACTTGCCTCATGTTATTGAGGAG 786

Qy        704 TTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAG 763

Db        787 GCTCCCTCCTATCCTGGTGTGGAGCATGTTGGTGGAGATATGTTTGTTAGTGTTCCAAAA 846

Qy        764 GCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAG 823

Db        847 GGAGATGCCATTTTCATGAAGTGGATATGTCATGATTGGAGCGATGAACACTGCTTGAAG 906

Qy        824 ATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTC 883

Db        907 TTTTTGAAGAAATGTTATGAAGCACTTCCA---------ACCAATGGGAAGGTGATCCTT 957

Qy        884 ATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTT 943

Db        958 GCTGAATGCATCCTCCCCGTGGCGCCAGACGCAAGCCTCCCCACTAAGGCAGTGGTCCAT 1017

Qy        944 ATGGATGTACACA---TGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGG 1000

Db       1018 ATTGATGTCATCATGTTGGCTCATAACCCAGGTGGGAAAGAGAGAACTGAGAAGGAGTTT 1077

Qy       1001 AAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCT 1049

Db       1078 GAGGCCTTGGCCAAGGGGGCTGGATTTGAAGGTTTCCGAGTAGTAGCCT 1126


RESULT 10
US-09-615-192A-107
; Sequence 107, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
;  APPLICANT:  Bloksberg, Leonard N.
;  APPLICANT:  Havukkala, Ilkka
;  TITLE OF INVENTION:  Materials and Methods for the
;  TITLE OF INVENTION:  Modification of Plant Lignin Content
;  FILE REFERENCE:  11000.1003c4U
;  CURRENT APPLICATION NUMBER: US/09/615,192A
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER:  US 08/975,316
;  PRIOR FILING DATE:  1997-11-21
;  PRIOR APPLICATION NUMBER:  US 08/713,000
;  PRIOR FILING DATE:  1996-09-11
;  PRIOR APPLICATION NUMBER:  US 09/169,789
;  PRIOR FILING DATE:  1998-10-09
;  NUMBER OF SEQ ID NOS:  405
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 107
;   LENGTH: 1630
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-615-192A-107

  Query Match             8.0%;  Score 92.6;  DB 4;  Length 1630;
  Best Local Similarity   50.8%;  Pred. No. 6.3e-16;
  Matches  280;  Conservative    0;  Mismatches  259;  Indels   12;  Gaps    2;

Qy        494 GACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAAC 553

Db        576 GACCCGCGATTCAACAAGATCTTTAACCGGGGAATGTCTGATCACTCCACCATTACTATG 635

Qy        554 TTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGT 613

Db        636 AAGAAGATACTGGAAACATACAAGGGCTTCGAGGGCCTCGAGACCGTGGTCGATGTCGGA 695

Qy        614 GGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATG 673

Db        696 GGCGGCACTGGGGCCGTGCTCAGCATGATCGTTGCCAAATACCCATCGATGAAAGGGATC 755

Qy        674 GTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTT 733

Db        756 AACTTCGACCTGCCTCACGTGATTGAAGACGCTCCACCCCTTCCTGGTGTCAAGCACGTC 815

Qy        734 GGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTA 793

Db        816 GGAGGCGACATGTTCGTCAGCGTTCCAAAGGGAGATGCCATTTTCATGAAGTGGATATGC 875

Qy        794 CATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCA 853

Db        876 CATGACTGGAGTGACGACCATTGCGCGAAGTTCCTCAAGAACTGCTACGATGCGCTTCC- 934

Qy        854 GGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGAT 913

Db        935 --------CAACAATGGAAAGGTGATCGTTGCAGAGTGCGTACTCCCTGTGTACCCAGAC 986
```

```
Qy        914 GAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACA---TGGCATGTATTATT 970
              | ||| ||  ||  |      | |||  || ||      ||   ||||    |
Db        987 ACGAGCCTAGCGACCAAGAATGTGATCCACATCGACTGCATCATGTTGGCCCACAACCCA 1046

Qy        971 AATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAA 1030
              || |||||||| | | |  | ||| |   ||  | |  ||  |  || || || ||
Db       1047 GGCGGGAAAGAGAGGACACAGAAGGAGTTCGAGGCATTGGCCAAAGGGGCCGGATTTCAG 1106

Qy       1031 AGCTACAAAAT 1041
              ||| | || |
Db       1107 GGCTTCCAAGT 1117


RESULT 11
US-09-615-192A-174
; Sequence 174, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
;  APPLICANT:  Bloksberg, Leonard N.
;  APPLICANT:  Havukkala, Ilkka
;  TITLE OF INVENTION:  Materials and Methods for the
;  TITLE OF INVENTION:  Modification of Plant Lignin Content
;  FILE REFERENCE:  11000.1003c4U
;  CURRENT APPLICATION NUMBER: US/09/615,192A
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER:  US 08/975,316
;  PRIOR FILING DATE:  1997-11-21
;  PRIOR APPLICATION NUMBER:  US 08/713,000
;  PRIOR FILING DATE:  1996-09-11
;  PRIOR APPLICATION NUMBER:  US 09/169,789
;  PRIOR FILING DATE:  1998-10-09
;  NUMBER OF SEQ ID NOS:  405
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 174
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-615-192A-174

  Query Match             7.7%;  Score 89.6;  DB 4;  Length 381;
  Best Local Similarity   58.8%;  Pred. No. 2.5e-15;
  Matches  174;  Conservative    0;  Mismatches  119;  Indels    3;  Gaps    1;

Qy        617 GGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTG 676
              | ||  | ||  || || |||       |  | || |||||  |  || ||||||   |
Db         10 GAAATGGAAACAACGGCCAAGCCATCGCGAAACGCCTTTCCGCATATGGAATGCACTATA 69

Qy        677 TTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGT 736
              || || | |||  |||| |||| |||||   ||  ||| | ||  ||| || ||||  |
Db         70 TTTGATCTTCCGCATGTGGTGGCCAATTTAGAAGTTAGCGAGAACGTGAGATGTGTTCCT 129

Qy        737 GGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACAT 796
              |||||||||||| | | ||| ||    || |||||| |  |  | |||| | |  | |||
Db        130 GGGGACATGTTTGAGTCCATACCACCAGCAGATGCAATAATATTGAAGTGGATACTCCAT 189

Qy        797 AATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTAT---TTCA 853
              ||||||  ||  | ||     ||||||| | |    || || || ||  |
Db        190 GATTGGAGCGATGAAGACGCTGTGAAGATACTGAAGCGATGCAAGGAGGCCTTAGGCAAG 249

Qy        854 GGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAA 909
              ||  |  |||| | |   || ||| || | || || |  ||||  || | |  ||
Db        250 GGCAAGGGCAAGAAACAGAAGGTAATTATAATTGACATGGTGATGGACAACACGAA 305


RESULT 12
US-09-615-192A-127
; Sequence 127, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
;  APPLICANT:  Bloksberg, Leonard N.
;  APPLICANT:  Havukkala, Ilkka
;  TITLE OF INVENTION:  Materials and Methods for the
;  TITLE OF INVENTION:  Modification of Plant Lignin Content
;  FILE REFERENCE:  11000.1003c4U
;  CURRENT APPLICATION NUMBER: US/09/615,192A
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER:  US 08/975,316
;  PRIOR FILING DATE:  1997-11-21
;  PRIOR APPLICATION NUMBER:  US 08/713,000
;  PRIOR FILING DATE:  1996-09-11
;  PRIOR APPLICATION NUMBER:  US 09/169,789
;  PRIOR FILING DATE:  1998-10-09
;  NUMBER OF SEQ ID NOS:  405
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 127
;   LENGTH: 661
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-615-192A-127

  Query Match             7.4%;  Score 86.2;  DB 4;  Length 661;
  Best Local Similarity   52.3%;  Pred. No. 2.6e-14;
  Matches  190;  Conservative    0;  Mismatches  173;  Indels    0;  Gaps    0;

Qy        502 ATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTT 561
              ||  ||||||   |||||   ||| ||| ||    |||||   |||||||       | |
Db         94 ATTCAACAAGATTTTCAACAGGGCTATGTCTGAGAATTCTACCATGTTGATGAACAAGAT 153

Qy        562 TAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAAC 621
              |  |||   |   ||  ||| |||   |  |   | ||||||| || || || | |
Db        154 TTTGGATACTTACGAGGGTTTTAAGGAGGTTCAGGAGTTGGTGGATGTGGGAGGAGGTAT 213

Qy        622 TGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGA 681
              |||    ||    ||  | ||          | |||  |  |   | | ||    || ||
Db        214 TGGGTCGACTCTCAATCTCATAGTGTCTAGGTATCCCCACATTTCAGGAATCAACTTCGA 273

Qy        682 ACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGA 741
              |  ||||  |||   ||    |       |      | ||| |  ||| ||||| ||
Db        274 CTTGTCCCATGTGCTGGCCGATGCTCCTCACTACCCAGCTGTGAAACATGTGGGTGGAGA 333

Qy        742 CATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTG 801
              ||||||| |  |  | || |  |   | ||  || |  | |||| | || | ||| ||||
Db        334 CATGTTTGATAGTGTACCAAGTGGCCAAGCTATTTTTATGAAGTGGATTCTGCATGATTG 393

Qy        802 GAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAG 861
              ||  ||  | |||||| |||| | ||  | ||||||  |  | ||  |  |||   | |
Db        394 GAGCGATGATCATTGCAGGAAGCTTTTGAAGAATTGTCACAAGGCGTTGCCAGAGAAGGG 453

Qy        862 CAA 864
              ||
Db        454 GAA 456


RESULT 13
US-08-975-316-51
; Sequence 51, Application US/08975316
; Patent No. 5952486
;  GENERAL INFORMATION:
;    APPLICANT:  BLOKSBERG, Leonard N., HAVUKKALA, Ilkka
;    APPLICANT:  and GRIERSON, Alastair W.
;    TITLE OF INVENTION:  MATERIALS AND METHODS FOR
;    TITLE OF INVENTION:  THE MODIFICATION OF PLANT LIGNIN CONTENT
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,316
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/713,000
;      FILING DATE:  September 11, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SLEATH, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000/1003C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  744 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-975-316-51

  Query Match             7.4%;  Score 86.2;  DB 2;  Length 744;
  Best Local Similarity   52.3%;  Pred. No. 2.7e-14;
  Matches  190;  Conservative    0;  Mismatches  173;  Indels    0;  Gaps    0;

Qy        502 ATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTT 561
              ||  ||||||   |||||   ||| ||| ||    |||||   |||||||         | |
Db        144 ATTCAACAAGATTTTCAACAGGGCTATGTCTGAGAATTCTACCATGTTGATGAACAAGAT 203

Qy        562 TAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAAC 621
              |   |||    |   ||  ||| |||   |  |      | ||||||| || || || |
Db        204 TTTGGATACTTACGAGGGTTTTAAGGAGGTTCAGGAGTTGGTGGATGTGGGAGGAGGTAT 263

Qy        622 TGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGA 681
              |||    ||   ||  | ||          | |||  |  |   | | ||    || ||
Db        264 TGGGTCGACTCTCAATCTCATAGTGTCTAGGTATCCCCACATTTCAGGAATCAACTTCGA 323

Qy        682 ACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGA 741
              |  ||||  |||   ||     |     | ||| |  ||| ||||| ||
Db        324 CTTGTCCCATGTGCTGGCCGATGCTCCTCACTACCCAGCTGTGAAACATGTGGGTGGAGA 383

Qy        742 CATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTG 801
              ||||||| |  |  | || |  |   | ||  || |  | |||| | || | ||| ||||
Db        384 CATGTTTGATAGTGTACCAAGTGGCCAAGCTATTTTTATGAAGTGGATTCTGCATGATTG 443

Qy        802 GAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAG 861
              ||  ||  |  |||||| |||| | ||  | |||||| |  | ||  |  |||   | |
Db        444 GAGCGATGATCATTGCAGGAAGCTTTTGAAGAATTGTCACAAGGCGTTGCCAGAGAAGGG 503

Qy        862 CAA 864
              ||
Db        504 GAA 506


RESULT 14
US-09-615-192A-51
; Sequence 51, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
;  APPLICANT:  Bloksberg, Leonard N.
;  APPLICANT:  Havukkala, Ilkka
;  TITLE OF INVENTION:  Materials and Methods for the
;  TITLE OF INVENTION:  Modification of Plant Lignin Content
;  FILE REFERENCE:  11000.1003c4U
;  CURRENT APPLICATION NUMBER: US/09/615,192A
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER:  US 08/975,316
;  PRIOR FILING DATE:  1997-11-21
;  PRIOR APPLICATION NUMBER:  US 08/713,000
;  PRIOR FILING DATE:  1996-09-11
;  PRIOR APPLICATION NUMBER:  US 09/169,789
;  PRIOR FILING DATE:  1998-10-09
;  NUMBER OF SEQ ID NOS:  405
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 51
;   LENGTH: 744
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-615-192A-51

  Query Match             7.4%;  Score 86.2;  DB 4;  Length 744;
  Best Local Similarity   52.3%;  Pred. No. 2.7e-14;
  Matches  190;  Conservative    0;  Mismatches  173;  Indels    0;  Gaps    0;

Qy        502 ATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTT 561
              ||  ||||||   |||||   ||| ||| ||    |||||   |||||||         | |
Db        144 ATTCAACAAGATTTTCAACAGGGCTATGTCTGAGAATTCTACCATGTTGATGAACAAGAT 203

Qy        562 TAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAAC 621
              |   |||    |   ||  ||| |||   |  |      | ||||||| || || || |
Db        204 TTTGGATACTTACGAGGGTTTTAAGGAGGTTCAGGAGTTGGTGGATGTGGGAGGAGGTAT 263

Qy        622 TGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGA 681
              |||    ||   ||  | ||          | |||  |  |   | | ||    || ||
Db        264 TGGGTCGACTCTCAATCTCATAGTGTCTAGGTATCCCCACATTTCAGGAATCAACTTCGA 323

Qy        682 ACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGA 741
              |  ||||  |||   ||     |     | ||| |  ||| ||||| ||
Db        324 CTTGTCCCATGTGCTGGCCGATGCTCCTCACTACCCAGCTGTGAAACATGTGGGTGGAGA 383

Qy        742 CATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTG 801
              ||||||| |  |  | || |  |   | ||  || |  | |||| | || | ||| ||||
Db        384 CATGTTTGATAGTGTACCAAGTGGCCAAGCTATTTTTATGAAGTGGATTCTGCATGATTG 443

Qy        802 GAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAG 861
              ||  ||  |  |||||| |||| | ||  | |||||| |  | ||  |  |||   | |
Db        444 GAGCGATGATCATTGCAGGAAGCTTTTGAAGAATTGTCACAAGGCGTTGCCAGAGAAGGG 503

Qy        862 CAA 864
              ||
Db        504 GAA 506


RESULT 15
US-08-204-288-6
; Sequence 6, Application US/08204288
; Patent No. 5959178
;  GENERAL INFORMATION:
;    APPLICANT:  VAN DOORSSELAERE, Jan
;    APPLICANT:  FRITIG, Bernard J.M.
;    APPLICANT:  INZE, Dirk G.
;    APPLICANT:  JOUANIN, Lise
;    APPLICANT:  KNIGHT, Mary E.
;    APPLICANT:  VAN MONTAGU, Marc
;    APPLICANT:  LEGRAND, Michel
;    TITLE OF INVENTION:  MODIFICATION OF LIGNIN SYNTHESIS IN
;    TITLE OF INVENTION:  PLANTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,288
;      FILING DATE:  10-MAR-1994
;      CLASSIFICATION:  800
```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9119279.9
; FILING DATE: 10-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01460
; FILING DATE: 09-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 206860/SEE36543/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-204-288-6

Query Match 7.1%; Score 82.6; DB 2; Length 1471;
Best Local Similarity 46.2%; Pred. No. 3.4e-13;
Matches 314; Conservative 0; Mismatches 359; Indels 6; Gaps 1;

```
Qy        425 GTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTT 484
              |||| ||| | ||   |  || |     |||  |   |    | |   |  || | |
Db        527 GTTCTTGAAGGAGGAGTTCCATTTGACAGGGTACACGGTGTTGTCCATGCATTTGAATAT 586

Qy        485 ATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAG 544
                ||    |||||   ||  ||  |    |||||  ||||||||     | |  |
Db        587 CCAAAATCGGACCCAAAATTCAATGATGTTTTCAACAAGGCAATGATCAATCACACAACT 646

Qy        545 ATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTG 604
              |  | |           |    ||     |   |  ||||||   ||  || |  | ||
Db        647 GTAGTCATGAAAAAAATACTTGAAAATTACAAAGGTTTTGAGAACCTTAAAACTTTGGTT 706

Qy        605 GATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTG 664
              |||||||| |||||   |||| | |     |||||  ||          |  || |   |
Db        707 GATGTTGGAGGTGGTCTTGGAGTTAACCTCAAGATGATTACATCTAAATACCCCACAATT 766

Qy        665 AAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTG 724
              ||  |||     || ||    ||| |||||||  || ||     |              ||
Db        767 AAGGGCACTAATTTTGATTTGCCACATGTTGTTCAACATGCCCCTTCCTATCCTGGGGTG 826

Qy        725 ACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAG 784
              |  |||||| || || |||||| || |  | ||  | |  |||||  || |  | |||
Db        827 GAACATGTTGGGGGAGATATGTTTGAAAGTGTTCCAGAAGGAGATGCTATTTTTATGAAG 886

Qy        785 TTGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAA 844
              | | || | ||| | |||| ||| |   |   | | ||| |  || | || ||  | |
Db        887 TGGATTCTTCATGACTGGAGTGATAGTCACAACCTCAAGTTGCTAAAGAACTGCTACAAG 946

Qy        845 GCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAA 904
              ||| |  |||   | | ||          || |  | |||     | |||| |   ||
Db        947 GCTCTACCAGACAATGGAAAGGTGATTGTTGTTGAGGCCATTTTACCAGTGAAACCAGAC 1006

Qy        905 AACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGT 964
              |       |||   |  |||   ||   |   ||     |       |    |||||
Db       1007 A------TTGACACCGCAGTGGTTGGCGTTTCGCAATGTGATTTGATCATGATGGCTCAA 1060

Qy        965 ATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGG 1024
              | |  |   || |||||| ||   ||||| || |     |   |  ||||| ||
Db       1061 AATCCTGGAGGCAAAGAGCGATCGGAAGAGGAGTTTCGAGCCTTGGCTACTGAAGCTGGA 1120

Qy       1025 TTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCT 1084
              ||| || ||   ||  ||      |       |  |||| |   | || ||  |||
Db       1121 TTCAAAGGCGTTAACTTAATATGTTGTGTCTGTAATTTTTGGGTCATGGAATTCTGCAAG 1180

Qy       1085 TGAATACTGACGCTGCAAT 1103
              |  ||    ||  | || |
```

Db 1181 TAGATTTCTACTGTACATT 1199

Search completed: November 10, 2003, 01:53:12
Job time : 105 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 10, 2003, 01:02:50 ; Search time 461 Seconds
(without alignments)
8029.308 Million cell updates/sec

Title: US-09-868-547-3
Perfect score: 1160
Sequence: 1 gtttgctattatggcttctt.....ataaaagttatctttgtgtt 1160

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2141354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	158.2	13.6	256	10	US-09-878-574-13958	Sequence 13958, A
2	142.8	12.3	266	10	US-09-878-574-5577	Sequence 5577, Ap
3	142.2	12.3	381	10	US-09-878-574-1726	Sequence 1726, Ap
4	120.2	10.4	1412	12	US-10-361-460-23	Sequence 23, Appl
5	111.8	9.6	377	10	US-09-878-574-3502	Sequence 3502, Ap
6	111.6	9.6	1315	12	US-10-361-460-24	Sequence 24, Appl
7	109.8	9.5	1503	10	US-09-947-027-5	Sequence 5, Appli
8	109.8	9.5	1503	13	US-10-091-009-5	Sequence 5, Appli
9	109	9.4	279	10	US-09-878-574-10987	Sequence 10987, A
10	108	9.3	1460	12	US-09-129-298-3	Sequence 3, Appli
11	107.6	9.3	272	10	US-09-878-574-13211	Sequence 13211, A
12	107.4	9.3	1418	12	US-09-129-298-4	Sequence 4, Appli
13	106.8	9.2	1306	12	US-10-361-460-25	Sequence 25, Appl
14	100.2	8.6	1149	10	US-09-938-842A-2582	Sequence 2582, Ap
15	94.6	8.2	1380	9	US-09-796-256A-5	Sequence 5, Appli
16	92.6	8.0	1630	15	US-10-174-693-107	Sequence 107, App
17	89.6	7.7	381	15	US-10-174-693-174	Sequence 174, App
18	86.2	7.4	661	15	US-10-174-693-127	Sequence 127, App
19	86.2	7.4	744	15	US-10-174-693-51	Sequence 51, Appl
20	84	7.2	221	10	US-09-878-574-7401	Sequence 7401, Ap
21	82.2	7.1	2096	14	US-10-137-036-60	Sequence 60, Appl
22	82.2	7.1	3070	14	US-10-137-036-113	Sequence 113, App
23	81.6	7.0	1494	14	US-10-253-971-5	Sequence 5, Appli
24	70.4	6.1	278	9	US-09-923-876-2383	Sequence 2383, Ap
25	63.8	5.5	279	10	US-09-878-574-9707	Sequence 9707, Ap
26	63.8	5.5	1440	12	US-10-289-757-43	Sequence 43, Appl
27	62.2	5.4	267	9	US-09-923-876-518	Sequence 518, App
28	61.6	5.3	398	10	US-09-878-574-48	Sequence 48, Appl
29	60.6	5.2	470	15	US-10-174-693-189	Sequence 189, App
30	59.2	5.1	687	14	US-10-213-473-25	Sequence 25, Appl
31	59	5.1	438	15	US-10-174-693-187	Sequence 187, App
32	57.8	5.0	1036	15	US-10-174-693-191	Sequence 191, App
33	57.6	5.0	736	14	US-10-213-473-27	Sequence 27, Appl
34	57	4.9	804	14	US-10-213-473-24	Sequence 24, Appl
35	56	4.8	1414	14	US-10-213-473-28	Sequence 28, Appl
36	54.6	4.7	1269	12	US-10-361-460-22	Sequence 22, Appl
37	54.4	4.7	1452	12	US-10-289-757-42	Sequence 42, Appl
38	54.2	4.7	682	15	US-10-174-693-192	Sequence 192, App
39	53.4	4.6	813	14	US-10-213-473-26	Sequence 26, Appl
40	52.6	4.5	254	9	US-09-923-876-1047	Sequence 1047, Ap
41	52.6	4.5	1455	12	US-10-289-757-150	Sequence 150, App
42	52.6	4.5	1457	12	US-10-289-757-41	Sequence 41, Appl
43	52	4.5	266	9	US-09-923-876-893	Sequence 893, App
44	51	4.4	1428	12	US-10-289-757-149	Sequence 149, App
45	51	4.4	1429	12	US-10-289-757-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-09-878-574-13958
; Sequence 13958, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 13958
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701068136H1
US-09-878-574-13958

Query Match 13.6%; Score 158.2; DB 10; Length 256;
Best Local Similarity 76.9%; Pred. No. 9.2e-33;
Matches 193; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```
Qy        515 TTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAAT 574
              |||||||||||| |||||  ||||||| |  ||  |  |||||||  | | | |||||| |
Db          6 TTCAATGAGGCTATGGCAAGTGATTCCCGAATAGTAGACTTGGCACTCAAAAATTGCACT 65

Qy        575 TGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCA 634
              |  || |||||||| || || ||||| ||||||||||||||||||||||||| ||| ||
Db         66 TCAGTTTTTGAGGGGCTAGATTCCATGGTGGATGTTGGTGGTGGAACTGGAACCACGGCC 125

Qy        635 AAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTT 694
              |  |||||||||| || |||||||| |||||||  | ||| | || | |||   ||||
Db        126 AGAATTATCTGTGACGCATTTCCTAAGTTGAAATGTGTTGTGCTTGACCTTCCTCCTGTT 185
```

```
Qy        695 GTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGC 754
Db        186 GTAGAGAACTTGACAGGGACCAATAATTTGAGTTTTGTTGGTGGTGACATGTTCAACTCT 245

Qy        755 ATCCCCAAGGC 765
Db        246 ATCCCTCAAGC 256

RESULT 2
US-09-878-574-5577
; Sequence 5577, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 5577
;   LENGTH: 266
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701097131H1
US-09-878-574-5577

  Query Match             12.3%;  Score 142.8;  DB 10;  Length 266;
  Best Local Similarity   72.1%;  Pred. No. 1.5e-28;
  Matches  186;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy        701 AATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCC 760
Db          2 AACTTGACAGGGACCAATAATTTGAGTTTTGTTGGTGGTGACATGTTCAACTCTATCCCT 61

Qy        761 AAGGCTGATGCAGTTCTGCTTAAGTTGGTTTACATAATTGGAATGACAACGATTGCATG 820
Db         62 CAAGCTGATGCAGTGCTACTAAAGTGGGTTTTACATAATTGGACCGACGAAAATTGCATA 121

Qy        821 AAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTT 880
Db        122 AAGATCCTGCAAAAGTGTAGAGATTCTATTTCAAGCAAAGGCAACAGTGGAAAAGTGATT 181

Qy        881 GTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTC 940
Db        182 ATCATAGATGCCGTAATAAATGAGAAGCTAGATGACCCGGATATGACACAAACAAAGCTT 241

Qy        941 CTTATGGATGTACACATG 958
Db        242 AGTTTGGACATTATTATG 259

RESULT 3
US-09-878-574-1726
; Sequence 1726, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 1726
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: LIB3028-034-Q1-B1-D6
US-09-878-574-1726

  Query Match             12.3%;  Score 142.2;  DB 10;  Length 381;
  Best Local Similarity   66.9%;  Pred. No. 2.6e-28;
  Matches  236;  Conservative    0;  Mismatches  108;  Indels    9;  Gaps    2;

Qy         21 CATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAAC 80
Db         25 CTTCAATGAATAACCAAAAAGAAATTGAGCTCTTTGAGGGCCAATCTCTTCTGTACATGC 84

Qy         81 ATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCG 140
Db         85 AGCTATATGGGCACCTAAGACCTATGTGTCTTAAGTGGGCTGTTCAACTAGGTATTCCAG 144

Qy        141 ACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTC 200
Db        145 ACATAAT------ACAGAACCATGCCAAACCCATTTCTCTTTCTGACTTGGTCTCTACTC 198

Qy        201 TACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACA 260
Db        199 TTCAAATTCCACCAGCTAACGCTGCTTTTGTGCAGCGGTTCATGCGCTTCTTGGCACACA 258

Qy        261 ATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGA---AGCATATGCTCTCACTG 317
Db        259 ATGGAATCTTTGAGATCCATGAGAGCCAAGAAGATCATGAACTAACATATGCTCTAACCC 318

Qy        318 CTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTT 370
Db        319 CTGCATCAAAGCTTCTTGTCAATAGTAGTGATCATTGTCTATCTCCAATGGTT 371

RESULT 4
US-10-361-460-23
; Sequence 23, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
;  APPLICANT: Helentjaris, Timothy G.
;  APPLICANT:  Bowen, Benjamin A.
;  APPLICANT:  Wang, Xun
;  TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
;  TITLE OF INVENTION:  Biosynthesis and Uses Thereof
;  FILE REFERENCE: 0709
;  CURRENT APPLICATION NUMBER: US/10/361,460
;  CURRENT FILING DATE:  2003-02-10
;  PRIOR APPLICATION NUMBER: 60/057,082
;  PRIOR FILING DATE: 1997-08-27
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
;   LENGTH: 1412
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (139)...(1263)
US-10-361-460-23

  Query Match             10.4%;  Score 120.2;  DB 12;  Length 1412;
  Best Local Similarity   54.0%;  Pred. No. 5.4e-22;
  Matches  292;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    2;

Qy        555 TGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTG 614
Db        734 TTGTGCTGAGGGAGTGCGGCGAGACGTTTCGTGGGATCGACTCGTTGGTTGACGTCGGTG 793

Qy        615 GTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGG 674
Db        794 GTGGCCATGGTGGCGCCGCCGCCACCATCGCCGCCGCCTTCCCCCACCTCAAGTGCAGCG 853
```

```
Qy        675 TGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTG 734
              || | || |  ||  | |||||  |     |          |||  ||   || || |
Db        854 TGCTTGACCTCCCGCACGTTGTCGCCGGTGCTCCGTCTGATGGCAACGTGCAGTTCGTCG 913

Qy        735 GTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTAC 794
                ||  | ||||||  |  | ||  ||    ||  ||| | ||  |||       | |||
Db        914 CAGGCAATATGTTTGAGAGTATTCCACCTGCAACCGCTGTTTTCCTCAAGAAAACTCTAC 973

Qy        795 ATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAG 854
              || | |||  |||| | || ||  | ||||||||  | |||||  ||  |||| || ||
Db        974 ATGACTGGGGTGACGATGAGTGTGTCAAGATATTGAAGAATTGCAAGCAAGCCATATCTC 1033

Qy        855 GTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATG 914
                    |  || || ||| |  || | ||   |||| ||   | | | |||  |
Db       1034 CACGGGATGCAGGTGGGAAGGTAATAATCTTGGA---TGTGGTAGTTGGATATAAACAGT 1090

Qy        915 AGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATG 974
                | |     | ||   | |  |||||  ||                    |||| |
Db       1091 CAAACATAAAGCATCAAGAGACACAAGTTATGTTTGATTTGTATATGATGGCGGTTAACG 1150

Qy        975 GAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCT 1034
              ||  ||| |  | ||  |||| ||||||||  ||||||  |||||  || ||| ||  ||
Db       1151 GAGTTGAGCGTGACGAGCAAGAGTGGAAGAAGATCTTCACTGAAGCTGGATTCAAAGACT 1210

Qy       1035 ACAAAATATCTCCCTTCACAGGATATTTGTC---TCTTATTGAGATCTATCCTTGAATAC 1091
              |||||||     ||| |||  ||  || | ||      | || ||| ||||||||||||| |
Db       1211 ACAAAATTCTACCCGTCATTGGTGATGTATCGGTCATCATCGAGGTCTATCCTTGAATGC 1270

Qy       1092 T 1092
              |
Db       1271 T 1271
```

RESULT 5
US-09-878-574-3502
; Sequence 3502, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3502
; LENGTH: 377
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-009-Q1-B1-H12
US-09-878-574-3502

Query Match 9.6%; Score 111.8; DB 10; Length 377;
Best Local Similarity 56.3%; Pred. No. 4.9e-20;
Matches 209; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

```
Qy        494 GACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAAC 553
              || ||    | |||||||   |||||  |||  ||||||  | | |||   ||    |
Db          7 GATCCAAGGTTTAACAAGGTTTTCAACAAGGGGATGGCTGATCACTCTACCATTACAATG 66

Qy        554 TTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGT 613
                |    ||   ||  | |    | ||||||||||||  ||||| | || |||||||||
Db         67 AAGAAAATTCTTGAGACCTACACAGGCTTTGAGGGACTTAAATCCCTGGTTGATGTTGGT 126

Qy        614 GGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATG 673
              || |||||||||     ||  || || ||          | ||| |   | ||  ||||
Db        127 GGAGGAACTGGAGCTGTAGTCAACATGATTGTCTCAAAGTATCCCACTATTAAGGGCATT 186
```

```
Qy        674 GTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTT 733
                || ||    ||  |||| | ||| ||    ||                ||  |  ||||
Db        187 AATTTTGATTTGCCCCATGTCATTGAAGATGCCCCATCTTATCCAGGAGTGGAACATGTT 246

Qy        734 GGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTA 793
              ||||| || ||||||    | | || || |||||||  || |  |||| | |||
Db        247 GGTGGAGATATGTTTGTCAGTGTTCCAAAAGCTGATGCTATTTTTATGAAGTGGATTTGC 306

Qy        794 CATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCA 853
              ||  |||||| |||  |  | ||| ||||| | ||  | || ||  | || ||  | ||
Db        307 CACGATTGGAGTGATGAGCACTGCTTGAAGTTTTTGAAGAACTGCTATGAGGCACTACCA 366

Qy        854 GGTGAAAGCAA 864
              | | |  | ||
Db        367 GATAATGGGAA 377
```

RESULT 6
US-10-361-460-24
; Sequence 24, Application US/10361460
; Publication No. US2003Q163839A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
; TITLE OF INVENTION: Biosynthesis and Uses Thereof
; FILE REFERENCE: 0709
; CURRENT APPLICATION NUMBER: US/10/361,460
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/057,082
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 1315
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)...(1160)
US-10-361-460-24

Query Match 9.6%; Score 111.6; DB 12; Length 1315;
Best Local Similarity 52.7%; Pred. No. 1.1e-19;
Matches 290; Conservative 0; Mismatches 254; Indels 6; Gaps 2;

```
Qy        546 TGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGG 605
              || ||    | | | | ||  | |||     | | |||   ||  |  | ||  | || |
Db        622 TGATGCCAATTGTGCTCAGGCAGTGCGGCGAGATGTTTCGTGGGATCAACTCATTGGTTG 681

Qy        606 ATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGA 665
              | || || |||||    |||   | | ||        |||   |  || || ||  | | |
Db        682 ACGTCGGCGGTGGGCATGGTGGCGCCGCCGCCGCCATCGCCGCTGCCTTCCCGCACGTCA 741

Qy        666 AATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGA 725
              | ||||  ||| | || |  ||  | |||||  |     |     ||         ||| ||
Db        742 AGTGCAGCGTGCTTGACCTCCCGCACGTTGTCGCCGGTGCTCCATCTGATGGCAACGTGC 801

Qy        726 CATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGT 785
              || || |  ||  | ||||||  |  | || ||    ||    || ||| | || |||
Db        802 AGTTCGTCGCAGGAAATATGTTTGAGAGTATTCCACCTGCAACCGCTGTTTTCCTCAAGA 861

Qy        786 TGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAG 845
              | ||||| | |||  |||| | || ||  | ||||||||  | ||||| ||  |||
Db        862 AAACTCTACATGACTGGGGTGACGATGAGTGTGTCAAGATATTGAAGAATTGCAAGCAAG 921

Qy        846 CTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAA 905
              | ||  |          |  ||||| ||| |  || | ||    ||| ||   | |
Db        922 CCATACCTCCACGGGATGCAGGTGGAAAGGTAATAATCTTGGA---CGTGGTAGTTGGAT 978

Qy        906 ACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTA 965
              | ||| |      | |      | ||   | |  |||||   |
```

```
Db        979 ATAAACAGTCAAACATAAAGCATCAAGAGACACAAGTTATGTTCGATTTGTATATGATGG 1038

Qy        966 TTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGT 1025

Db       1039 CCGTTAACGGAGTTGAGCGTGACGAGCAAGAGTGGAAGAAGATCTTCGCCGAAGCCGGAT 1098

Qy       1026 TCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTC---TTATTGAGATCTATC 1082

Db       1099 TCAAAGACTACAAAATTCTACCCGTCATTGGTGACGTGTCGGTCATCATCGAGGTCTATC 1158

Qy       1083 CTTGAATACT 1092

Db       1159 CTTGAATGCT 1168


RESULT 7
US-09-947-027-5
; Sequence 5, Application US/09947027
; Patent No. US20020124281A1
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent Lee C.
;  APPLICANT:  Li, Laigeng
;  TITLE OF INVENTION: METHODS FOR SIMULTANEOUS CONTROL OF LIGNIN CONTENT AND COMPOSITIO
;  TITLE OF INVENTION:  CELLULOSE CONTENT IN PLANTS
;  FILE REFERENCE: 066040-9718
;  CURRENT APPLICATION NUMBER: US/09/947,027
;  CURRENT FILING DATE:  2001-09-05
;  PRIOR APPLICATION NUMBER: 60/230,086
;  PRIOR FILING DATE: 2000-09-05
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 1503
;   TYPE: DNA
;   ORGANISM: aspen populus tremuloides
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: AldOMT; GenBank accession number:  X62096
US-09-947-027-5

  Query Match             9.5%;  Score 109.8;  DB 10;  Length 1503;
  Best Local Similarity   49.6%;  Pred. No. 3.8e-19;
  Matches  313;  Conservative    0;  Mismatches  312;  Indels    6;  Gaps    1;

Qy        448 ATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAA 507

Db        523 ATTTAACAAGGCCTATGGGATGACTGCATTTGAATATCATGGCACGGATCCAAGATTCAA 582

Qy        508 CAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGA 567

Db        583 CAAGGTCTTCAACAAGGGAATGTCTGACCACTCTACCATTACCATGAAGAAGATTCTTGA 642

Qy        568 TTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAAT 627

Db        643 GACCTACAAAGGCTTTGAAGGCCTCACGTCCTTGGTGGATGTTGGTGGTGGGACTGGAGC 702

Qy        628 CACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCC 687

Db        703 CGTCGTTAACACCATCGTCTCTAAATACCCTTCAATCAAGGGCATTAACTTCGATCTGCC 762

Qy        688 AAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTT 747

Db        763 CCACGTCATTGAGGATGCCCCATCTTATCCCGGAGTGGAGCATGTTGGTGGCGACATGTT 822

Qy        748 TAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGA 807

Db        823 TGTTAGTGTGCCCAAAGCAGATGCCGTTTTCATGAAGTGGATATGCCATGATTGGAGCGA 882

Qy        808 CAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAAC 867

Db        883 CGCCCACTGCTTAAAATTCTTGAAGAATTGCTATGACGCGTTGCCGGAAAACGGCAAGGT 942

Qy        868 AGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTAC 927

Db        943 GATACTTGTTGAGTGCATTCTTCCCGTGGCTCCTGACACAAGCCTTGCCACCAAGGG--- 999

Qy        928 TGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAA 987

Db       1000 ---AGTCGTGCACGTTGATGTCATCATGCTGGCGCACAACCCCGGTGGGAAAGAGAGGAC 1056

Qy        988 AGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCC 1047

Db       1057 CGAGAAGGAATTTGAGGGCTTAGCTAAGGGAGCTGGCTTCCAAGGTTTTGAAGTAATGTG 1116

Qy       1048 CTTCACAGGATATTTGTCTCTTATTGAGATC 1078

Db       1117 CTGTGCATTCAACACACATGTCATTGAATTC 1147


RESULT 8
US-10-091-009-5
; Sequence 5, Application US/10091009
; Publication No. US20020138870A1
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent Lee C.
;  APPLICANT:  Li, Laigeng
;  TITLE OF INVENTION: METHODS FOR SIMULTANEOUS CONTROL OF LIGNIN CONTENT AND
;  TITLE OF INVENTION:  COMPOSITION, AND
;  TITLE OF INVENTION:  CELLULOSE CONTENT IN PLANTS
;  FILE REFERENCE: 066040-9718
;  CURRENT APPLICATION NUMBER: US/10/091,009
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: 09/947,027
;  PRIOR FILING DATE: 2001-09-05
;  PRIOR APPLICATION NUMBER: 60/230,086
;  PRIOR FILING DATE: 2000-09-05
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 1503
;   TYPE: DNA
;   ORGANISM: aspen populus tremuloides
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: AldOMT; GenBank accession number:  X62096
US-10-091-009-5

  Query Match             9.5%;  Score 109.8;  DB 13;  Length 1503;
  Best Local Similarity   49.6%;  Pred. No. 3.8e-19;
  Matches  313;  Conservative    0;  Mismatches  312;  Indels    6;  Gaps    1;

Qy        448 ATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAA 507

Db        523 ATTTAACAAGGCCTATGGGATGACTGCATTTGAATATCATGGCACGGATCCAAGATTCAA 582

Qy        508 CAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGA 567

Db        583 CAAGGTCTTCAACAAGGGAATGTCTGACCACTCTACCATTACCATGAAGAAGATTCTTGA 642

Qy        568 TTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAAT 627

Db        643 GACCTACAAAGGCTTTGAAGGCCTCACGTCCTTGGTGGATGTTGGTGGTGGGACTGGAGC 702

Qy        628 CACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCC 687

Db        703 CGTCGTTAACACCATCGTCTCTAAATACCCTTCAATCAAGGGCATTAACTTCGATCTGCC 762

Qy        688 AAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTT 747

Db        763 CCACGTCATTGAGGATGCCCCATCTTATCCCGGAGTGGAGCATGTTGGTGGCGACATGTT 822

Qy        748 TAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGA 807

Db        823 TGTTAGTGTGCCCAAAGCAGATGCCGTTTTCATGAAGTGGATATGCCATGATTGGAGCGA 882

Qy        808 CAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAAC 867
```

```
Db        883 CGCCCACTGCTTAAAATTCTTGAAGAATTGCTATGACGCGTTGCCGGAAAACGGCAAGGT 942

Qy        868 AGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTAC 927

Db        943 GATACTTGTTGAGTGCATTCTTCCCGTGGCTCCTGACACAAGCCTTGCCACCAAGGG--- 999

Qy        928 TGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAA 987

Db       1000 ---AGTCGTGCACGTTGATGTCATCATGCTGGCGCACAACCCCGGTGGGAAAGAGAGGAC 1056

Qy        988 AGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCC 1047

Db       1057 CGAGAAGGAATTTGAGGGCTTAGCTAAGGGAGCTGGCTTCCAAGGTTTTGAAGTAATGTG 1116

Qy       1048 CTTCACAGGATATTTGTCTCTTATTGAGATC 1078

Db       1117 CTGTGCATTCAACACACATGTCATTGAATTC 1147
```

RESULT 9
US-09-878-574-10987
; Sequence 10987, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 10987
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701063901H1
US-09-878-574-10987

Query Match 9.4%; Score 109; DB 10; Length 279;
Best Local Similarity 64.6%; Pred. No. 2.4e-19;
Matches 179; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

```
Qy        158 AGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACT 217

Db          2 AACCATGCCAAACCCATTTCTCTTTCTGACTTGGTCTCTACTCTTCAAATTCCACCAGCT 61

Qy        218 AAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTCTTTGAGATA 277

Db         62 AACGCTGCCTTTGTGCAGCGGTTCATGCGCTTCTTGGCACACAATGGAATCTTTGAGATC 121

Qy        278 GTAAGAATCCATGACAACATAGA---AGCATATGCTCTCACTGCTGCTTCAGAGTTACTT 334

Db        122 CATGAGCGCCAAGAAGATCATGAACTAACATATGCTCTAACCCCTGCATCAAAGCTTCTT 181

Qy        335 GTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGT 394

Db        182 GTCAATAGTAGTACTCATTGTCTATCTCCAATGGTTCTAGCGTTTACCGATCCACTTCGG 241

Qy        395 CAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATG 431

Db        242 AACGTTAAATACCATCACTTGGGGGCATGGCTTCGTG 278
```

RESULT 10
US-09-129-298-3
; Sequence 3, Application US/09129298
; Publication No. US20030196218A1
; GENERAL INFORMATION:
; APPLICANT: Arntzen, Charles
; APPLICANT: Kipp, Peter B.
; APPLICANT: Kumar, Ramesh
; APPLICANT: May, Gregory D.
; TITLE OF INVENTION: The Use of Mixed Duplex Oligonucleotides
; TITLE OF INVENTION: to Effect Localized Genetic Changes in Plants
; FILE REFERENCE: 7991-023-999
; CURRENT APPLICATION NUMBER: US/09/129,298
; CURRENT FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/054,386
; EARLIER FILING DATE: 1997-08-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(1178)
US-09-129-298-3

Query Match 9.3%; Score 108; DB 12; Length 1460;
Best Local Similarity 52.6%; Pred. No. 1.1e-18;
Matches 259; Conservative 0; Mismatches 230; Indels 3; Gaps 1;

```
Qy        404 TGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTA 463

Db        501 TGGTACCACTTAAAAGATGCAGTACTAGATGGTGGAATCCCATTCAACAAGGCCTATGGA 560

Qy        464 GGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAG 523

Db        561 ATGACAGCATTTGAGTACCATGGC---ACAGATCCAAGATTCAACAAAGTTTTCAACCGT 617

Qy        524 GCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTT 583

Db        618 GGAATGTCTGATCACTCCACTATGTCAATGAAAAAGATTCTTGAGGACTACAAAGGATTT 677

Qy        584 GAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATC 643

Db        678 GAAGGCCTAAATTCCATTGTTGATGTTGGTGGTGGAACTGGCGCTACTGTTAACATGATT 737

Qy        644 TGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAAT 703

Db        738 GTCTCCAAACATCCCTCTATTAAGGGTATTAACTTTGATTTACCACATGTTATTGGAGAT 797

Qy        704 TTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAG 763

Db        798 GCTCCAGCTTACCCTGGTGTCGAGCACGTTGGTGGCGACATGTTTGCCAGTGTGCCAAAA 857

Qy        764 GCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAG 823

Db        858 GCAGATGCCATTTTCATGAAGTGGATTTGTCATGATTGGAGCGACGAGCATTGCCTAAAA 917

Qy        824 ATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTC 883

Db        918 TTCTTGAAGAATTGCTATGAAGCACTACCTGCAAATGGGAAGGTGATAATAGCGGAGTGC 977

Qy        884 ATAGATACTGTG 895

Db        978 ATACTTCCAGAG 989
```

RESULT 11
US-09-878-574-13211
; Sequence 13211, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants

```
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 13211
;   LENGTH: 272
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701066773H1
US-09-878-574-13211

  Query Match             9.3%;  Score 107.6;  DB 10;  Length 272;
  Best Local Similarity   66.8%;  Pred. No. 5.6e-19;
  Matches  171;  Conservative    0;  Mismatches   79;  Indels    6;  Gaps    1;

Qy         21 CATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAAC 80
Db         11 CTTCAATGAATAACCAAAAAGAAATTGAGCTCTTTGAGGGCCAATCTCTTCTGTACATGC 70

Qy         81 ATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCG 140
Db         71 AGCTATATGGGCACCTAAGACCTATGTGTCTTAAGTGGGCTGTTCAACTAGGTATTCCAG 130

Qy        141 ACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTC 200
Db        131 ACATAAT------ACAGACCCATGCCAAACCCATTTCTCTTTCTGACTTGGTCTCTACTC 184

Qy        201 TACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACA 260
Db        185 TTCAAATTCCACCAGCTAACGCTGCCTTTGTGCAGCGGTTCATGCGCTTCTTGGCACACA 244

Qy        261 ATGGATTCTTTGAGAT 276
Db        245 ATGGAATCTTTGAGAT 260


RESULT 12
US-09-129-298-4
; Sequence 4, Application US/09129298
; Publication No. US20030196218A1
; GENERAL INFORMATION:
;  APPLICANT: Arntzen, Charles
;  APPLICANT: Kipp, Peter B.
;  APPLICANT: Kumar, Ramesh
;  APPLICANT: May, Gregory D.
;  TITLE OF INVENTION: The Use of Mixed Duplex Oligonucleotides
;  TITLE OF INVENTION: to Effect Localized Genetic Changes in Plants
;  FILE REFERENCE: 7991-023-999
;  CURRENT APPLICATION NUMBER: US/09/129,298
;  CURRENT FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/054,386
;  EARLIER FILING DATE: 1997-08-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 1418
;   TYPE: DNA
;   ORGANISM: Nicotiana tabacum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (59)...(1153)
US-09-129-298-4

  Query Match             9.3%;  Score 107.4;  DB 12;  Length 1418;
  Best Local Similarity   52.9%;  Pred. No. 1.6e-18;
  Matches  231;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy        459 CCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCA 518
Db        528 CCTATGGAATGACAGCATTTGAGTACCATGGCACAGATCCAAGATTCAACAAAGTGTTCA 587

Qy        519 ATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGG 578
Db        588 ACCGTGGAATGTCTGATCACTCCACTATGTCAATGAAGAAGATTCTTGAGGACTACAAAG 647

Qy        579 TCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGA 638
Db        648 GATTTGAAGGCCTAAATTCCATTGTTGATGTTGGTGGTGGAACGGGTGCTACTGTTAACA 707

Qy        639 TTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGG 698
Db        708 TGATTGTCTCTAAATATCCCTCTATTAAGGGCATTAACTTTGATTTGCCACATGTAATTG 767

Qy        699 AAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCC 758
Db        768 GAGATGCTCCAACTTACCCCGGTGTCGAGCACGTTGGTGGCGACATGTTTGCTAGTGTGC 827

Qy        759 CCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCA 818
Db        828 CAAAAGCAGATGCCATTTTCATGAAGTGGATTTGTCATGATTGGAGCGATGAGCATTGCC 887

Qy        819 TGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAG 878
Db        888 TAAAATTCTTGAAGAATTGCTATGAAGCACTACCTGCAAATGGGAAGGTGATAATTGCAG 947

Qy        879 TTGTCATAGATACTGTG 895
Db        948 AGTGCATACTTCCAGAG 964


RESULT 13
US-10-361-460-25
; Sequence 25, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
;  APPLICANT: Helentjaris, Timothy G.
;  APPLICANT:  Bowen, Benjamin A.
;  APPLICANT:  Wang, Xun
;  TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
;  TITLE OF INVENTION:  Biosynthesis and Uses Thereof
;  FILE REFERENCE: 0709
;  CURRENT APPLICATION NUMBER: US/10/361,460
;  CURRENT FILING DATE:  2003-02-10
;  PRIOR APPLICATION NUMBER: 60/057,082
;  PRIOR FILING DATE: 1997-08-27
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 1306
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (57)...(1154)
US-10-361-460-25

  Query Match             9.2%;  Score 106.8;  DB 12;  Length 1306;
  Best Local Similarity   51.3%;  Pred. No. 2.3e-18;
  Matches  273;  Conservative    0;  Mismatches  256;  Indels    3;  Gaps    1;

Qy        577 GGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAA 636
Db        650 GGTGTTCCGTGGGATCGACTCGTTGGTCGACGTCGGCGGTGGGCACGGTGGCGCCGCCGC 709

Qy        637 GATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGT 696
Db        710 CACCATCGCCGCCGCATTCCCGCACGTCAAGTGCAGCGTGCTTGACCTCCCGCACGTTGT 769

Qy        697 GGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCAT 756
Db        770 CGCCGGTGCTCCATCCGATGCCTGCGTGCAGTTCGTTGCGGGCAATATGTTCCACAGTAT 829

Qy        757 CCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTG 816
```

```
Db        830 TCCACCTGCAACCGCCGTTTTCTTCAAGACAACTCTATGTGACTGGGGTGACGACGAGTG 889

Qy        817 CATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGT 876

Db        890 CATCAAGATATTGAAGAATTGCAAGCAAGCCATATCTCCACGGGATGAGGGTGGGAAGGT 949

Qy        877 AGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAA 936

Db        950 GATAATCATGGACGTGGTAGTCGGGTATGGGCAGTCAAACATGAAGCGCCTAGAGACACA 1009

Qy        937 GCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGA 996

Db       1010 AGTTATGTTTGATTTGGTTATGATGGCGG---TCAATGGAGTCGAGCGCGACGAGCAAGA 1066

Qy        997 TTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGG 1056

Db       1067 GTGGAAGGAGATGTTCATTGAAGCTGGATTCAAAGACTACAAAATCCGACCAGTAGCTGG 1126

Qy       1057 ATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGCTGCAATATTCC 1108

Db       1127 CCTCATGTCGGTCATCGAGGTCTATCCATGAATTCTTTGTGAACAAAAGGCC 1178
```

RESULT 14
US-09-938-842A-2582
; Sequence 2582, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2582
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2582

```
  Query Match             8.6%;  Score 100.2;  DB 10;  Length 1149;
  Best Local Similarity   56.4%;  Pred. No. 1.3e-16;
  Matches  208;  Conservative    0;  Mismatches  158;  Indels    3;  Gaps    1;

Qy        516 TCAATGAGGCAATGGCTTGTGATTC---TCAGATGTTGAACTTGGCGTTTAGAGATTGCA 572

Db        554 TCAATGAGGCCATGGCTTGTGATGCAAGGCGCGTGGTGCCACGTGTAGCCGGAGCTTGTC 613

Qy        573 ATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAG 632

Db        614 ACGGCTTGTTTGATGGCGTGACTACGATGGTTGACGTAGGAGGTGGTACGGGAGAGACGA 673

Qy        633 CAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATG 692

Db        674 TGGGGATGCTTGTGAAGGAGTTTCCTTGGATCAAAGGATTTAACTTTGATCTTCCTCATG 733

Qy        693 TTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAAT 752

Db        734 TCATTGAAGTTGCTGAAGTCTTGGACGGTGTTGAGAATGTTGAGGGCGATATGTTTGATT 793

Qy        753 GCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACG 812

Db        794 CTATTCCGGCTTGCGACGCCATTTTCATCAAGTGGGTGTTACACGATTGGGGAGACAAAG 853

Qy        813 ATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAA 872

Db        854 ATTGCATAAAGATATTGAAGAATTGCAAAGAAGCGGTCCCTCCAAATATCGGAAAAGTGT 913

Qy        873 AAGTAGTTG 881

Db        914 TGATAGTGG 922
```

RESULT 15
US-09-796-256A-5
; Sequence 5, Application US/09796256A
; Patent No. US20020078477A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Vincent L
; APPLICANT: Carraway, Daniel T
; APPLICANT: Smeltzer, Richard H
; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
; FILE REFERENCE: 50617/c-3532.0
; CURRENT APPLICATION NUMBER: US/09/796,256A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US60/033381
; PRIOR FILING DATE: 1996-12-16
; PRIOR APPLICATION NUMBER: 08/991677
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Liquidambar styraciflua
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(1170)
US-09-796-256A-5

```
  Query Match             8.2%;  Score 94.6;  DB 9;  Length 1380;
  Best Local Similarity   50.1%;  Pred. No. 4.9e-15;
  Matches  325;  Conservative    0;  Mismatches  309;  Indels   15;  Gaps    3;

Qy        404 TGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTA 463

Db        490 TGGTACCACTTGACCGAGGCAGTTCTTGAAGGTGGAATTCCATTTAACAAGGCCTATGGA 549

Qy        464 GGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAG 523

Db        550 ATGACAGCATTTGAGTACCATGGC---ACCGATCCCAGATTCAACACAGTTTTCAACAAT 606

Qy        524 GCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTT 583

Db        607 GGAATGTCCAATCATTCGACCATTACCATGAAGAAAATCCTTGAGACTTACAAAGGGTTC 666

Qy        584 GAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATC 643

Db        667 GAGGGACTTGGATCTGTGGTTGATGTTGGTGGTGGCACTGGTGCCCACCTTAACATGATT 726

Qy        644 TGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAAT 703

Db        727 ATCGCTAAATACCCCATGATCAAGGGCATTAACTTCGACTTGCCTCATGTTATTGAGGAG 786

Qy        704 TTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAG 763

Db        787 GCTCCCTCCTATCCTGGTGTGGAGCATGTTGGTGGAGATATGTTTGTTAGTGTTCCAAAA 846

Qy        764 GCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAG 823

Db        847 GGAGATGCCATTTTCATGAAGTGGATATGTCATGATTGGAGCGATGAACACTGCTTGAAG 906

Qy        824 ATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTC 883

Db        907 TTTTGAAGAAATGTTATGAAGCACTTCCA---------ACCAATGGGAAGGTGATCCTT 957
```

```
Qy        884 ATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTT 943
              ||      |  |   ||        ||| |   |||     ||| |   |  | ||| |
Db        958 GCTGAATGCATCCTCCCCGTGGCGCCAGACGCAAGCCTCCCCACTAAGGCAGTGGTCCAT 1017

Qy        944 ATGGATGTACACA---TGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGG 1000
              || |||||   ||   ||||   ||       ||| ||||||||||  ||  | || |
Db       1018 ATTGATGTCATCATGTTGGCTCATAACCCAGGTGGGAAAGAGAGAACTGAGAAGGAGTTT 1077

Qy       1001 AAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCT 1049
              ||    |   || ||  || || ||  || | | |  | ||    |||
Db       1078 GAGGCCTTGGCCAAGGGGGCTGGATTTGAAGGTTTCCGAGTAGTAGCCT 1126
```

Search completed: November 10, 2003, 03:17:19
Job time : 463 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 9, 2003, 23:36:19 ; Search time 2914 Seconds
(without alignments)
9675.083 Million cell updates/sec

Title: US-09-868-547-3
Perfect score: 1160
Sequence: 1 gtttgctattatggcttctt.....ataaaagttatctttgtgtt 1160

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	660.4	56.9	697	9	AW350290	AW350290 GM210008A
	2	588	50.7	589	12	BM523459	BM523459 sam84a09.
	3	557.8	48.1	561	13	BQ452916	BQ452916 sao93b08.
	4	548	47.2	573	12	BI245401	BI245401 FiS1C9 Fo
	5	493	42.5	494	9	AI507864	AI507864 sa88b04.y
	6	456.2	39.3	852	10	BG588377	BG588377 EST490186
	7	442.6	38.2	822	10	BG645368	BG645368 EST506987
	8	439.2	37.9	569	12	BM188258	BM188258 saj98f01.
	9	437.4	37.7	606	12	BM177364	BM177364 saj80d09.
	10	417.8	36.0	788	13	BQ165419	BQ165419 EST611276
c	11	415.8	35.8	693	14	CA921971	CA921971 EST639689
	12	413	35.6	546	12	BM178815	BM178815 saj59h07.
	13	410	35.3	826	13	BQ165418	BQ165418 EST611275
	14	405	34.9	625	9	AW733746	AW733746 sk77f10.y
c	15	405	34.9	775	13	BQ165420	BQ165420 EST611277
	16	397.2	34.2	774	9	AW559420	AW559420 EST314468
	17	390.6	33.7	742	9	AW267810	AW267810 EST305938
	18	387	33.4	408	13	BU765089	BU765089 sar75g12.
	19	385.8	33.3	503	10	BG511336	BG511336 sad16e12.
	20	383.2	33.0	420	10	BE021456	BE021456 sm49c06.y
	21	379.6	32.7	718	13	BU091459	BU091459 st73e04.y
	22	375.8	32.4	512	10	BG509454	BG509454 sad14a06.
c	23	364	31.4	551	9	AW309379	AW309379 sf16f05.x
	24	353.8	30.5	519	10	BG509170	BG509170 sac87b12.
	25	348.8	30.1	668	13	BQ156708	BQ156708 NF096C11I
	26	348.8	30.1	746	14	CA783152	CA783152 sat20d01.
	27	338	29.1	563	12	BM528698	BM528698 sak67e02.
	28	330.8	28.5	639	10	BE999705	BE999705 EST431428
c	29	330.2	28.5	763	9	AW348354	AW348354 GM210002A
	30	329.6	28.4	643	10	BF648180	BF648180 NF045A09E
	31	327.2	28.2	519	10	BE943072	BE943072 EST422651
	32	327	28.2	327	9	AI441334	AI441334 sa55d11.y
	33	320.8	27.7	592	13	BU577747	BU577747 sar92a10.
	34	318	27.4	634	10	BF650939	BF650939 NF102C07E
	35	310.4	26.8	779	14	CB891286	CB891286 EST648255
c	36	309.6	26.7	496	13	BU549771	BU549771 GM880023B
	37	295	25.4	615	10	BG448331	BG448331 NF070A11E
	38	290	25.0	305	13	BQ627748	BQ627748 sap35g06.
	39	287.4	24.8	577	9	AW559532	AW559532 EST314580
	40	285.4	24.6	623	9	AW584374	AW584374 N210406e
	41	284.4	24.5	577	9	AW585248	AW585248 N211531e
	42	281	24.2	541	12	BI417247	BI417247 LjNEST49a
	43	280.8	24.2	546	9	AI507849	AI507849 sa87h11.y
	44	279.4	24.1	600	10	BF637366	BF637366 NF027F03P
	45	269.2	23.2	591	13	BU764527	BU764527 sas02g11.

ALIGNMENTS

RESULT 1
AW350290/c
LOCUS AW350290 697 bp mRNA linear EST 04-OCT-2000
DEFINITION GM210008A10D3 Gm-r1021 Glycine max cDNA clone Gm-r1021-2837 3', mRNA sequence.
ACCESSION AW350290
VERSION AW350290.1 GI:6848000
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 697)
AUTHORS Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V., Erpelding,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.
TITLE A Functional Genomics Program for Soybean (NSF 9872565)
JOURNAL Unpublished
COMMENT Other_ESTs: AI507864
Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147

Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888)919-3324 or (314) 427-3324 or contact:clones@genomesystems.com or info@genomesystems.com web site:www.genomesystems.com
Seq primer: 5'-TTTTTTTTTTTTTTTTTTTT(A/C/G)-3'.

FEATURES Location/Qualifiers
source 1. .697
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="Gm-r1021-2837"
/tissue_type="root"
/lab_host="XL10-Gold"
/clone_lib="Gm-r1021"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; Library Gm-r1021 is a sequence-driven, reracked set of the original library Gm-c1004 which was prepared from root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. The Gm-c1004 library was constructed by Dr. Paul Keim & Virginia H. Coryell, Department of Biology, Box5640, Northern Arizona University, Flagstaff, AZ 86011 , email: paul.keim@nau.edu, virginia.coryell@nau.edu. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Computational Biology Centers, University of Minnesota, http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html . Reracking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html."

BASE COUNT 198 a 157 c 97 g 221 t 24 others
ORIGIN

Query Match 56.9%; Score 660.4; DB 9; Length 697;
Best Local Similarity 96.3%; Pred. No. 3.3e-133;
Matches 672; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

```
Qy        445 AGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATA 504
              ||||||||| | ||||||| |||| | | ||||||||||||||||||||||||||||||
Db        697 AGTATTTGANNNNNCNTTAGGAA-NNNNNNNNGGGACNNNNNCNNTNAAGACCCTGCATA 639

Qy        505 TAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAG 564
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        638 TAACAAGTCATTCAATGAGGCAATGGCTNGTGATTCTCAGATGTTGAACTTGGCGTTTAG 579

Qy        565 AGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGG 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 AGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGG 519

Qy        625 AATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACG 684
              |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db        518 AATCACAGCAAAGATTATCTGTGAGGCTTTTCTTNAGCTGAAATGCATGGTGTTGGAACG 459

Qy        685 TCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACAT 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 TCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACAT 399

Qy        745 GTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAA 804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAA 339

Qy        805 TGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAA 864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 TGACAACGATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAA 279
```

```
Qy        865 AACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGT 924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 AACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGT 219

Qy        925 TACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAG 984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 TACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAG 159

Qy        985 AAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATC 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 AAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATC 99

Qy       1045 TCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGCTGCAATA 1104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 TCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGCTGCAATA 39

Qy       1105 TTCCATTTAGTAGTTAATTTGCATGTTATCAATAAAAT 1142
              |||||||||||||||||||||||||||||||||||||
Db         38 TTCCATTTAGTAGTTAATTTGCATGTTATCAATAAAAT 1
```

RESULT 2
BM523459
LOCUS BM523459 589 bp mRNA linear EST 19-FEB-2002
DEFINITION sam84a09.y2 Gm-c1087 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1087-426 5' similar to TR:O22555 O22555 O-METHYTRANSPERASE. [2] TR:O22308 ;, mRNA sequence.
ACCESSION BM523459
VERSION BM523459.1 GI:18723299
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 589)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 438.

FEATURES Location/Qualifiers
source 1. .589
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1087-426"
/tissue_type="Roots"
/lab_host="DH10B"
/clone_lib="Gm-c1087"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The mRNA was prepared using polyatract mRNA system from PROMEGA. The cDNA was prepared using the STRATAGENE kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site

(5'GAGAGAGAGAGAGAGAGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTT). EcoRI adapters (5'OH-AATTCGGCACGAG and 3'GCCGTGCTCp) were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction sites of the pBluescript vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH10B by electroporation. Library is in LB medium with ampicillin and glycerol 8%. Average insert size: 800 bp."

BASE COUNT 164 a 106 c 143 g 175 t 1 others
ORIGIN

Query Match 50.7%; Score 588; DB 12; Length 589;
Best Local Similarity 99.8%; Pred. No. 1.8e-117;
Matches 588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        193 GTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCT 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCT 60

Qy        253 AGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCT 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCT 120

Qy        313 CACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGA 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGA 180

Qy        373 GTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGA 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGA 240

Qy        433 GGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAA 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAA 300

Qy        493 AGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAA 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAA 360

Qy        553 CTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGG 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGG 420

Qy        613 TGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCAT 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCAT 480

Qy        673 GGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGT 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGT 540

Qy        733 TGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTT 781
              |||||||||||||||| |||||||||||||||||||||||||||||||
Db        541 TGGTGGGGACATGTTTANATGCATCCCCAAGGCTGATGCAGTTCTGCTT 589
```

RESULT 3
BQ452916
LOCUS BQ452916 561 bp mRNA linear EST 29-MAY-2002
DEFINITION sao93b08.y1 Gm-c1081 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1081-3159 5' similar to TR:O22309 O22309 7-O-METHYLTRANSFERASE. ;, mRNA sequence.
ACCESSION BQ452916
VERSION BQ452916.1 GI:21256028
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 561)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 449.
FEATURES Location/Qualifiers
source 1. .561
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1081-3159"
/tissue_type="Roots of 7 day old 'Bragg' seedlings"
/dev_stage="7 days old"
/lab_host="DH10B"
/clone_lib="Gm-c1081"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of 7 day old 'Bragg' seedlings that were mock-infected 48 hours prior to harvest. Dr. Gary Stacey generously donated the tissue. The roots were flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first- strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constucted in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University."
BASE COUNT 185 a 79 c 132 g 165 t
ORIGIN

Query Match 48.1%; Score 557.8; DB 13; Length 561;
Best Local Similarity 99.6%; Pred. No. 6.6e-111;
Matches 559; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        565 AGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGG 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGG 60
```

```
Qy        625 AATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACG 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACG 120

Qy        685 TCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACAT 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACAT 180

Qy        745 GTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAA 804
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTTAAATGCTTTCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAA 240

Qy        805 TGACAACGATTGCATGAAGATATTAGAAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAA 864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGACAACGATTGCATGAAGATATTAGAAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAA 300

Qy        865 AACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGT 924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACAGGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGT 360

Qy        925 TACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAG 984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACTGAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAG 420

Qy        985 AAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATC 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAAGAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATC 480

Qy       1045 TCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGCTGCAATA 1104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCCTTCACAGGATATTTGTCTCTTATTGAGATCTATCCTTGAATACTGACGCTGCAATA 540

Qy       1105 TTCCATTTAGTAGTTAATTTG 1125
              |||||||||||||||||||||
Db        541 TTCCATTTAGTAGTTAATTTG 561
```

```
RESULT 4
BI245401
LOCUS       BI245401                573 bp    mRNA    linear   EST 30-DEC-2001
DEFINITION  FiS1C9 Forrest infected Subtraction Library Glycine max cDNA, mRNA
            sequence.
ACCESSION   BI245401
VERSION     BI245401.1  GI:17998777
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Iqbal,M.J. and Lightfoot,D.A.
  TITLE     Differentially expressed mRNA sequences identified by subtraction
            hybridization from soybean, Glycine max (L.) Merr. var. Forrest,
            roots in response to Fusarium solani f. sp. glycinae inoculation
  JOURNAL   Unpublished
COMMENT     Contact: Iqbal MJ and Lightfoot, DA
            Center of Excellence in Soybean Research, Teaching and Outreach
            Southern Illinois University at Carbondale
            Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
            USA
            Tel: 618 453-3121
            Fax: 618 453-7457
            Email: mjiqbal@siu.edu, ga4082@siu.edu
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /cultivar="Forrest"
                     /db_xref="taxon:3847"
                     /dev_stage="14 days after inoculation"
                     /clone_lib="Forrest infected Subtraction Library"
                     /note="Organ: Root; Vector: pGEM; Forrest roots were
                     inoculated with Fusarium solani f. sp. glycinae and
                     samples were collected after 14 days of inoculation. Total
                     RNA was used  for cDNA synthesis using SMART PCR cDNA
                     synthesis Kit (CLONTECH, CA). CLONTECH PCR-Select
                     Subtraction system was used to enrich rare  transcripts
                     expressed in Forrest roots under inoculation by F. solani
                     and cloned in pGEM vector. ESTs were sequenced using M13
                     universal  forward or reverse primer."
BASE COUNT      169 a    116 c    119 g    169 t
ORIGIN
```

```
  Query Match           47.2%;  Score 548;  DB 12;  Length 573;
  Best Local Similarity  100.0%;  Pred. No. 9e-109;
  Matches  548;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTTTGCTATTATGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 GTTTGCTATTATGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGG 85

Qy         61 TCAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 TCAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGAT 145

Qy        121 GGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTT 205

Qy        181 TTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCT 265

Qy        241 CATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGA 325

Qy        301 AGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGC 385

Qy        361 TCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 TCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAG 445

Qy        421 GTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGA 505

Qy        481 CTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTC 565

Qy        541 TCAGATGT 548
              ||||||||
Db        566 TCAGATGT 573
```

```
RESULT 5
AI507864
LOCUS       AI507864                494 bp    mRNA    linear   EST 28-NOV-2001
DEFINITION  sa88b04.y1 Gm-c1004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1004-6344 5' similar to TR:O22308 O22308 7-O-METHYLTRANSFERASE.
            [3] TR:O22309 TR:O22555 ;, mRNA sequence.
ACCESSION   AI507864
VERSION     AI507864.1  GI:4405715
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
```

```
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Insert Length: 1307   Std Error: 0.00
            Seq primer: -40RP from Gibco
            High quality sequence stop: 393
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1004-6344"
                     /tissue_type="root"
                     /lab_host="XL10-Gold"
                     /clone_lib="Gm-c1004"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; Root cDNA. The mRNA was isolated from entire roots
                     of 8 day old 'Williams' seedlings which were propagated on
                     paper towels with distilled water. Stratagene's  cDNA
                     Synthesis Kit (catalog #200401) was used to synthesize the
                     cDNA. First- strand synthesis was performed with 5-methyl
                     dCTP, hence the ligated cDNA is hemimethylated.
                     Stratagene's first-strand synthesis primer was used
                     [GAGAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)-18]. After
                     second-strand synthesis, the cDNA ends were 'polished'
                     with clone Pfu DNA polymerase, ligated to EcoRI adapters,
                     and phosphorylated. The XhoI site within the first-strand
                     synthesis primer was restricted by digestion with XhoI;
                     all XhoI sites in the  cDNA would be protected by their
                     hemimethylated status. The cDNA constructs were
                     size-fractionated with a 500bp cutoff, using GibcoBRL Life
                     Technologies' cDNA Size Fractionation column. The column
                     eluent was then ligated into  Stratagene's pBluescript II
                     XR Predigested vector (pBluescript II SK(+) that had been
                     digested with EcoRI and XhoI, and phosphorylated). Both
                     the white  and blue colonies appear to contain recombinant
                     plasmids with cDNA inserts. Blue colonies 9n=15) have been
                     sequenced, and possess putative cDNA inserts. This library
                     was constructed by Dr. Paul Keim & Virginia H. Coryell,
                     Department of Biology, Box5640, Northern Arizona
                     University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.
                     Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax:
                     520-523-7500, email: paul.keim@nau.edu,
                     virginia.coryell@nau.edu"
BASE COUNT      144 a    103 c    101 g    145 t      1 others
ORIGIN

  Query Match             42.5%;  Score 493;  DB 9;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 8.1e-97;
  Matches  493;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         12 TGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCT 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCT 60

Qy         72 TGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTG 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTG 120

Qy        132 ACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGG 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGG 180

Qy        192 TGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATC 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATC 240

Qy        252 TAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTC 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGAAGCATATGCTC 300

Qy        312 TCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTG 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTG 360

Qy        372 AGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATG 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATG 420

Qy        432 AGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATA 491
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        421 AGGAAGATCTCACAGTATTTGAGGTCTNCTTAGGAACACCTTTCTGGGACTTTATCAATA 480

Qy        492 AAGACCCTGCATAT 505
              ||||||||||||||
Db        481 AAGACCCTGCATAT 494


RESULT 6
BG588377
LOCUS       BG588377                852 bp    mRNA    linear   EST 12-APR-2001
DEFINITION  EST490186 MHRP- Medicago truncatula cDNA clone pMHRP-53C11, mRNA
            sequence.
ACCESSION   BG588377
VERSION     BG588377.1  GI:13606517
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 852)
  AUTHORS   Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
            and Fraser,C.M.
  TITLE     ESTs from phosphate-starved roots of Medicago truncatula, 2001
  JOURNAL   Unpublished
COMMENT     Contact: Harrison M.J.
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73401
            Tel: 580-223-5810
            Fax: 580-221-7380
            Email: mjharrison@noble.org
            The Samuel Roberts Noble Foundation: N387128e TIGR sequence name:
            MTHAY18TK More information is available at: http://www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .852
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pMHRP-53C11"
                     /tissue_type="roots"
                     /dev_stage="phosphate-starved"
                     /lab_host="XLOLR"
                     /clone_lib="MHRP-"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
```

```
              XhoI; At the trifoliate stage, M. truncatula plants were
              transplanted to phosphate-free sand and grown for a
              further 30 days. During this period, they were fertilized
              twice weekly with 1/2 Hoaglands  solutions containing 20uM
              potassium phosphate. cDNA was prepared from polyA+
              enriched RNA. The cDNA was directionally ligated into the
              Unizap XR vector from Stratagene and packaged using
              Gigapack III Gold packaging extracts. Plasmids containing
              cDNA inserts  were excised from the recombinant lambda-Zap
              phage using Ex-assist  helper phage and propagated in
              XLOLR cells."
BASE COUNT      268 a    121 c    189 g    274 t
ORIGIN

  Query Match             39.3%;  Score 456.2;  DB 10;  Length 852;
  Best Local Similarity   73.8%;  Pred. No. 9.2e-89;
  Matches  613;  Conservative    0;  Mismatches  203;  Indels   15;  Gaps    2;

Qy        106 GTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGG 165
Db         28 GTCTCTTAAATGGTCAGTTGAAATGGACATACCAACTATTATCTAC------AATCATGG 81

Qy        166 CCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCG 225
Db         82 CAAACCAATTACTCTTTCAAACTTAGTTTCAATTCTTCAAATTCCATCAACCAAAGTCGA 141

Qy        226 TCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAAT 285
Db        142 TAACGTGCAGCGTCTCATGCGTCTCCTTGCACACAATGGATTCTTTGAGATAGTAACGAA 201

Qy        286 CCATG---------ACAACATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGT 336
Db        202 TCAAGAATTGGAAAATGAAGAAGAAGCTTATGCTCTCACAGTTACTTCAGAGCTTCTTGT 261

Qy        337 CAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCA 396
Db        262 TAAAGGGACTGAACTTTGTTTAGCTCCAATGGTTGAGTGTGTTCTTGATCCTAAATTTAA 321

Qy        397 AGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGT 456
Db        322 TGTTTCATTTCATAATTTCAAGAAGTGGATTTATGAGGAAGATCTCACCCTCTTTGCTAC 381

Qy        457 CTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATT 516
Db        382 CTCTTTAGGATGTGATTTATGGGAGTTTCTTAACAAAAATCCTGAATATAATAGATTGTT 441

Qy        517 CAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTG 576
Db        442 CAATGATGCTATGGCTAGTGATTCCAAAATGATTAATTTGGCATTGAAAGATTGCAATTT 501

Qy        577 GGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAA 636
Db        502 TGTGTTTGAGGGGTTGGATTCAATTGTGGATGTTGGTGGTGGAAATGGAACAACTGGTAA 561

Qy        637 GATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGT 696
Db        562 GATTATTTGTGAGACATATCCTAAGTTGAGATGTGTTGTTTTTGATCGTCCAAAGGTTGT 621

Qy        697 GGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCAT 756
Db        622 AGAGAATTTATGTGGAAGTAACAATTTGACATATGTCGGTGGGGACATGTTCATATCTGT 681

Qy        757 CCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTG 816
Db        682 TCCTAAAGCTGATGCAGTTCTGCTTAAGTGGATTTTACATGATTGGACGGATAGGGATTG 741

Qy        817 CATGAAGATATTAGAAAATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGT 876
Db        742 CATAAAGATATTGAAGAAATGTAAAGAAGCTGTTACAACTGAGGGGAAAAGAGGAAAAGT 801

Qy        877 AGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTAC 927
Db        802 GATTGTGATAGACATGGTAATAAATGAAAAGGAAGGAAGAGACGAACTTAC 852
```

```
RESULT 7
BG645368
LOCUS       BG645368                  822 bp    mRNA    linear   EST 24-APR-2001
DEFINITION  EST506987 KV3 Medicago truncatula cDNA clone pKV3-46C11 5' end,
            mRNA sequence.
ACCESSION   BG645368
VERSION     BG645368.1  GI:13780480
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   VandenBosch,K., Endre,G., Hur,J., Beremand,P., Town,C.D., Bowman
            ,C.L., Craven,M.B., Cho,J. and Fraser,C.M.
  TITLE     ESTs from roots of Medicago truncatula 72 h after Rhizobium
            inoculation, 2001
  JOURNAL   Unpublished
COMMENT     Contact: VandenBosch K
            Department of Plant Biology
            University of Minnesota
            220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
            Tel: 612 624 2755
            Fax: 612 625 1738
            Email: kvandenb@cbs.umn.edu
            M393775e TIGR sequence name: MTECA18TK More information is
            available at: www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1..822
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /cultivar="genotype A17"
                     /db_xref="taxon:3880"
                     /clone="pKV3-46C11"
                     /tissue_type="Seedling roots"
                     /dev_stage="3 days post-inoculation with Sinorhizobium
                     meliloti"
                     /lab_host="E. coli strain XLOLR"
                     /clone_lib="KV3"
                     /note="Vector: pBluescript SK -; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                     was directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts.  Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in XLOLR cells."
BASE COUNT      250 a    124 c    169 g    279 t
ORIGIN

  Query Match             38.2%;  Score 442.6;  DB 10;  Length 822;
  Best Local Similarity   74.2%;  Pred. No. 8.3e-86;
  Matches  605;  Conservative    0;  Mismatches  194;  Indels   16;  Gaps    3;

Qy         18 CTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACA 77
Db         15 CTTCTTCAATAAATAACCGTAAAGCAAGTGAAATTTTCAAAGCACAAGCTCTATTATACA 74

Qy         78 AACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATAC 137
Db         75 AAAACATGTATGCTTTCATAGATTCCATGTCTCTTAAATGGTCAGTTGAAATGGACATAC 134

Qy        138 CCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAA 197
Db        135 CAACTATTATCTAC------AATCATGGCAAACCAATTACTCTTTCAAACTTAGTTTCAA 188

Qy        198 TTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCAC 257
Db        189 TTCTTCAAATTCCATCAACCAAAGTCGATAACGTGCAGCGTCTCATGCGTCTCCTTGCAC 248
```

```
Qy        258 ACAATGGATTCTTTGAGATAGTAAGAATCCATG---------ACAACATAGAAGCATATG 308
              ||||||||||||||||||||||||||  |  || |         |   |||||| ||||
Db        249 ACAATGGATTCTTTGAGATAGTAACGAATCAAGAATTGGAAAATGAAGAAGAAGCTTATG 308

Qy        309 CTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGG 368
              ||||||| | | |||||||| | ||||| ||| | | ||| ||| ||||||||||||||||
Db        309 CTCTCACAGTTACTTCAGAGCTTCTTGTTAAAGGGACTGAACTTTGTTTAGCTCCAATGG 368

Qy        369 TTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTC 428
              |||||| | ||||||| || || | | | | |||   |  | || |||| |||| ||
Db        369 TTGAGTGTGTTCTTGATCCTAAATTTAATGTTTCATTTCATAATTTCAAGAAGTGGATTT 428

Qy        429 ATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCA 488
              |||||||||||||||  | ||||    ||| ||||||    ||| ||||| ||| | |
Db        429 ATGAGGAAGATCTCACCCTCTTTGCTACCTCTTTAGGATGTGATTTATGGGAGTTTCTTA 488

Qy        489 ATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGT 548
              | ||| | |||| |||||| | |  ||||||| || |||||| |||||||  | |||
Db        489 ACAAAAATCCTGAATATAATAGATTGTTCAATGATGCTATGGCTAGTGATTCCAAAATGA 548

Qy        549 TGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATG 608
              | || ||||| || | ||||||||||  || ||||||||  |||| || ||||||||||
Db        549 TTAATTTGGCATTGAAAGATTGCAATTTTGTGTTTGAGGGGTTGGATTCAATTGTGGATG 608

Qy        609 TTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAAT 668
              |||||||||||| |||||  || |  |||||||| ||||||  | | |||||||| ||| ||
Db        609 TTGGTGGTGGAAATGGAACAACTGGTAAGATTATTTGTGAGACATATCCTAAGTTGAGAT 668

Qy        669 GCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACAT 728
              |  | || || || |||||||| ||||| || ||||| |  ||||| ||||||||||||||
Db        669 GTGTTGTTTTTGATCGTCCAAAGGTTGTAGAGAATTTATGTGGAAGTAACAATTTGACAT 728

Qy        729 TTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGG 788
              |||  |||||||||||||| | ||   | || ||  ||||||||||||||||||||||| |
Db        729 ATGTCGGTGGGGACATGTTCATATCTGTTCCTAA-ACTGATGCAGTTCTGCTTAAGTGGA 787

Qy        789 TTTTACATAATTGGAATGACAACGATTGCATGAAG 823
              |||||||| ||||||  || |  |||||||| |||
Db        788 TTTTACATGATTGGACGGATAGGGATTGCATAAAG 822
```

RESULT 8
BM188258

LOCUS	BM188258 569 bp mRNA linear EST 11-DEC-2001
DEFINITION	saj98f01.y1 Gm-c1074 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1074-4849 5' similar to TR:O22555 O22555 O-METHYTRANSFERASE. [2] TR:O22308 ;, mRNA sequence.
ACCESSION	BM188258
VERSION	BM188258.1 GI:17519216
KEYWORDS	EST.
SOURCE	Glycine max (soybean)
ORGANISM	Glycine max Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE	1 (bases 1 to 569)
AUTHORS	Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com Seq primer: -40RP from Gibco High quality sequence stop: 425.

```
FEATURES             Location/Qualifiers
     source          1. .569
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="SOYBEAN CLONE ID: Gm-c1074-4849"
                     /tissue_type="seedlings induced for HR (hypersensitive
                     response)"
                     /dev_stage="9-11 day old"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1074"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from mRNA isolated
                     from 9-11 day old seedlings that were induced for HR
                     (hypersensitive response) by vacuum infiltrating plant
                     tissue with Pseudomonas syringae pv. glycinea carrying the
                     avrB gene (Genetics 141:1597-1604). Plant tissue (expanded
                     unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36,
                     and 53 hrs after inoculation and their mRNA pooled equally
                     for cDNA construction. The library was prepared using the
                     Stratagene pBluescript II SK(+) library construction kit.
                     Complementary DNA was synthesized from mRNA using a primer
                     consisting of a poly(dT) sequence with an XhoI restriction
                     site. EcoRI adaptors were ligated to the blunt-ended cDNA
                     fragments followed by XhoI digestion. The cDNA insert is
                     protected from XhoI digestion via methylation during first
                     strand synthesis. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into E.coli ElectroMax DH10B host cells. Plant
                     care, inoculations, and library construction were
                     performed by Steve Clough (Lila Vodkin lab, University of
                     Illinois)."
BASE COUNT      160 a    101 c    137 g    171 t
ORIGIN
```

Query Match 37.9%; Score 439.2; DB 12; Length 569;
Best Local Similarity 86.7%; Pred. No. 4.2e-85;
Matches 496; Conservative 0; Mismatches 73; Indels 3; Gaps 1;

```
Qy        213 CAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTCTTTG 272
              ||  |||||  | |||| ||||| || |||||||| || || ||||||||||||||||||
Db          1 CAGCTAAAGTTAGTCAAGTGCAAAGTCTCATGCGCTACCTGGCACACAATGGATTCTTTG 60

Qy        273 AGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTAC 332
              | | |||||  |||||||| ||   |||||||||||||||||||||||||||||||| |||
Db         61 AAAGAGTAACAATCCATGAAAA---AGAAGCATATGCTCTCACTGCTGCTTCAGAGCTAC 117

Qy        333 TTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATT 392
              | |||||||||||||||||||||||||||||||||||||||||||||||| |||||||  ||| |
Db        118 TAGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATATTCTTGACACAACTA 177

Qy        393 GTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTG 452
              |  |||| |||   | |||||||||| |||||||||||||||||||||||||| ||||||
Db        178 TTTCAGGTTCATTTCATCAGTTGAAGAAGTGGGTTCATGAGGAAGATCTCACACTATTTG 237

Qy        453 AGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGT 512
              || | ||||||||| ||| ||| ||||||||| | |||| | ||||||||||||||||||
Db        238 AGATTTCCTTAGGATCACATTTGTGGGACTTTCTTAATAGAAACCCTGCATATAACAAGT 297

Qy        513 CATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCA 572
              |||||||||||||||||||| |||||||||||||||||||||||||||| ||||||||||
Db        298 CATTCAATGAGGCAATGGCTAGTGATTCTCAGATGTTGAACTTGGCGTTGAGAGATTGCA 357

Qy        573 ATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAG 632
              | | ||| |||||||||||||||||||||||||||||||||||||||||||||| ||| |
Db        358 AGTTGGTGTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAGCCACTG 417
```

```
Qy        633 CAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATG 692
              | ||||||||||||||| || |||||| |  |||||||||| ||||| || |||||  | |
Db        418 CCAAGATTATCTGTGAAGCATTTCCTGACTTGAAATGCATTGTGTTTGACCGTCCGCAGG 477

Qy        693 TTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAAT 752
              ||||||| || ||||||||||||||||||||||||| |||||| |||||||||| ||||
Db        478 TTGTGGAGAACTTGTCAGGAAGCAACAATTTGACATATGTTGGAGGGGACATGTTCAAAT 537

Qy        753 GCATCCCCAAGGCTGATGCAGTTCTGCTTAAG 784
                || ||||||||   |||||||||  |||||
Db        538 CTATTCCCAAGGCCTGTGCAGTTCTATTTAAG 569
```

RESULT 9
BM177364
LOCUS BM177364 606 bp mRNA linear EST 06-DEC-2001
DEFINITION saj80d09.y1 Gm-c1074 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1074-2873 5' similar to TR:O22555 O22555 O-METHYTRANSFERASE. [2] TR:O22308 ;, mRNA sequence.
ACCESSION BM177364
VERSION BM177364.1 GI:17400582
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 606)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 442.
FEATURES Location/Qualifiers
source 1. .606
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1074-2873"
/tissue_type="seedlings induced for HR (hypersensitive response)"
/dev_stage="9-11 day old"
/lab_host="DH10B"
/clone_lib="Gm-c1074"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."
BASE COUNT 178 a 122 c 119 g 187 t
ORIGIN

Query Match 37.7%; Score 437.4; DB 12; Length 606;
Best Local Similarity 85.6%; Pred. No. 1.1e-84;
Matches 525; Conservative 0; Mismatches 76; Indels 12; Gaps 3;

```
Qy          1 GTTTGCTATTATGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGG 60
              |||||| |||||||||||||||   || ||||||||||||| | ||||||||| ||||||||
Db          6 GTTTGCAATTATGGCTTCTTC---AATCAATGGCCGTAATGAAAGTGAGATCTTTCAAGG 62

Qy         61 TCAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGAT 120
              |||| |||||||||||||||| ||| || |||||||||| |||| |||||| ||||| ||
Db         63 TCAAACTCTCTTGTACAAACACTTGTATGCCTTCATAGACTCTATGTGTCTCAAATGCAT 122

Qy        121 GGTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTT 180
              |||||||||  ||||| ||||||||||||      | ||||||||||||||||||| |
Db        123 TGTTGAGCTTGGAATACCAGACATAATCCAC------AACCATGGCCAACCCATTACTCT 176

Qy        181 TTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCT 240
              | | ||||||||||||||||||| | |||||| |||||  | |||| ||||| || ||
Db        177 TCCTGAGTTGGTGTCAATTCTACAAATTCCACCAGCTAAAGTTAGTCAAGTGCAAAGTCT 236

Qy        241 CATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGACAACATAGA 300
              |||||| || || |||||||||||||||||||| | |||||||||||| || ||   |||
Db        237 CATGCGCTACCTGGCACACAATGGATTCTTTGAAAGAGTAAGAATCCACGAAAA---AGA 293

Qy        301 AGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTTAGTTTAGC 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        294 AGCATATGCTCTCACTGCTGCTTCAGAGCTACTTGTCAAAAGCAGTGAGCTTAGTTTAGC 353

Qy        361 TCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAGTTGAAGAG 420
              ||||||| |||||| | ||||||| |||| ||    | | |||   | || ||||| |
Db        354 TCCAATGATTGAGTTTGTTCTTGACCCAACTTTGTCAAATTCATTCCATCAATTGAAAAA 413

Qy        421 GTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCTTTCTGGGA 480
              ||||||| ||||| |||||||||| |||||||  |||||||||| ||| ||| |||||
Db        414 GTGGGTTTATGAGAAAGATCTCACATTATTTGACATCTCCTTAGGATCACATTTGTGGGA 473

Qy        481 CTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTGATTC 540
              |||| | |||||| |||||||| ||||||||||||||||||||||||||||| |||||||
Db        474 CTTTCTTAATAAAAACCCTGCACATAACAAGTCATTCAATGAGGCAATGGCTAGTGATTC 533

Qy        541 TCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAATCCAT 600
              ||||||| |||||||||||| |||||||||||||||||||| | |||||||||| |||
Db        534 TCAGATGATGAACTTGGCGTTGAGAGATTGCAATTGGGTCTTTCAAGGACTGGAATTCAT 593

Qy        601 TGTGGATGTTGGT 613
              |||||||||||||
Db        594 TGTGGATGTTGGT 606
```

RESULT 10
BQ165419
LOCUS BQ165419 788 bp mRNA linear EST 25-APR-2002
DEFINITION EST611276 KVKC Medicago truncatula cDNA clone pKVKC-9A5, mRNA sequence.
ACCESSION BQ165419
VERSION BQ165419.1 GI:20307786
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 788)
AUTHORS VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S., Utterback,T., Cheung,F. and Fraser,C.M.
TITLE The Medicago truncatula 'kiloclone' set; ESTs selected and re-arrayed from various libraries
JOURNAL Unpublished
COMMENT Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cbs.umn.edu
TIGR sequence name: MTNAI05TK Alias Clone name:DSIR-19L16 More information is available at: www.medicago.org
Seq primer: SKmod (CTA gAA CTA gtg gAT CC).

FEATURES Location/Qualifiers
source 1. .788
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pKVKC-9A5"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XLOLR"
/clone_lib="KVKC"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

BASE COUNT 227 a 119 c 166 g 276 t
ORIGIN

Query Match 36.0%; Score 417.8; DB 13; Length 788;
Best Local Similarity 73.7%; Pred. No. 2e-80;
Matches 577; Conservative 0; Mismatches 192; Indels 14; Gaps 3;

```
Qy          33 GCCGTAAAGCAAGTGAGATTTTTCAAGGTCAAGCTCTCTTGTACAAACATTTGCTTGGCT 92
               | |  ||| | ||||| |||||  | |  |||||| | || ||||||||  |  ||| ||
Db          18 GGCAGAAAACCAGTGAAATTTTCAAGGCACAAGCTTTATTATACAAACACATATTTGCCT 77

Qy          93 TCATAGATTCTAAGTGTCTAAAATGGATGGTTGAGCTTGACATACCCGACATAATCCACA 152
               |||||||||| | || ||| ||||||   |||||  |  |||||||  | || || |||
Db          78 TCATAGATTCCATGTCTCTTAAATGGGCTGTTGAAATGAACATACCAAATATTATTCAC- 136

Qy         153 GCCATAGCCATGGCCAACCCATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTCCCAC 212
                 |  |||||| |||| ||| |  ||||| | || || |||||||| ||||| |||
Db         137 -----AATCATGGCAAACCAATTTCACTTTCAAATTTAGTTTCAATTCTTCAAGTTCCAT 191

Qy         213 CAACTAAAACTCGTCAGGTGCAGAGCCTCATGCGTTATCTAGCACACAATGGATTCTTTG 272
               |  | ||||   || | |||| | | ||||||||||| || || ||||| ||||||||||
Db         192 CGTCGAAAATAGGTAACGTGCGGCGTCTCATGCGTTACCTCGCGCACAACGGATTCTTTG 251

Qy         273 AGATAGTAAGAATCCATGACAACATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTAC 332
               |||||  |  ||   | ||      |||  | |||||||||||| |||||||||| | |
Db         252 AGATAATTACAAAAGAAGA------AGAGTCTTATGCTCTCACTGTTGCTTCAGAGCTTC 305

Qy         333 TTGTCAAAAGCAGTGAGCTTAGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATT 392
               ||||  |  |  ||||| ||| ||||||| |||||||||||| | ||||||| |||| |
Db         306 TTGTTAGAGGTAGTGATCTTTGTTTAGCACCAATGGTTGAGTGTGTTCTTGATCCAACTC 365

Qy         393 GTCAAGGTGCATGGAACCAGTTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTG 452
                  |||  |  |   |  || ||||||  ||| || |||||||||||| ||| | ||||
Db         366 TTTCGGGTTCGTATCATGAGCTGAAGAAATGGATTTATGAGGAAGATCTTACACTCTTTG 425

Qy         453 AGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGT 512
               ||  | |||||  |   ||| |||| ||| |  |||| | |||| ||||||  |  |
Db         426 GTGTTACTTTAGGGTCTGGTTTTTGGGATTTTCTTGATAAAAATCCTGAATATAATAGAT 485

Qy         513 CATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCA 572
               |||| ||||| |||||||| |||||||  ||  | ||||||| || ||||||||
Db         486 CATTTAATGATGCAATGGCTAGTGATTCTAAATTGATAAACTTGGCATTGAGAGATTGTG 545

Qy         573 ATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAG 632
               |||  || ||||| ||| ||||||| |||||||||||||||||||| ||||| ||| |
Db         546 ATTTTGTGTTTGATGGATTGGAATCAATTGTGGATGTTGGTGGTGGAAATGGAACCACTG 605

Qy         633 CAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATG 692
               || ||||| |||||| |||||||||| |||||||  || ||||| ||  | ||| | |
Db         606 GTAAAATTATTTGTGAGACTTTTCCTAAGTTGAAATGTATTGTGTTTGATAGGCCACAAG 665

Qy         693 TTGTGGAAAATTTGTCAGGAAGCAACAATTTGACAT--TTGTTGGTGGGGACATGTTTAA 750
               |||| || || || || |||||||| |||||||  |  | |  ||||||||||| |
Db         666 TTGTAGAGAACTTATCTGGAAGCAATAATTTGACTTTATGTTGGTGGGGGACATGTTCAC 725

Qy         751 ATGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAA 810
               ||  || || || |||||||||  | |||||||||   ||||||||||||||| ||| ||
Db         726 ATCTATTCCTAATGCTGATGCACGTTTGCTTAAGTATATTTTACATAATTGGACTGATAA 785

Qy         811 CGA 813
               |||
Db         786 GGA 788
```

RESULT 11
CA921971/c
LOCUS CA921971 693 bp mRNA linear EST 09-MAY-2003
DEFINITION EST639689 MTUS Medicago truncatula cDNA clone MTUS-46H12, mRNA sequence.
ACCESSION CA921971
VERSION CA921971.1 GI:27408901
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 693)
AUTHORS VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S., Utterback,T., Cheung,F. and Fraser,C.M.
TITLE The Medicago truncatula 6K unigene set: cDNA clones selected and re-arrayed from various libraries
JOURNAL Unpublished
COMMENT Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cbs.umn.edu
Alias Clone pDSIR-19L16
TIGR sequence name: MTUBT96TV
More information is available at: www.medicago.org
Seq primer: (gtA AtA CgA CtC ACt AtA ggg C).

FEATURES Location/Qualifiers
source 1. .693
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="MTUS-46H12"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XLOLR"
/clone_lib="MTUS"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA

was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

BASE COUNT 240 a 147 c 79 g 227 t
ORIGIN

Query Match 35.8%; Score 415.8; DB 14; Length 693;
Best Local Similarity 77.5%; Pred. No. 5.4e-80;
Matches 517; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

```
Qy        474 TCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTT 533
              | | ||| ||| |  ||||| | |||| |||||| |  ||||| ||||| |||||||||
Db        693 TTTTGGATTTTCTTGATAAAAATCCTGAATATAATAGATCATTTAATGATGCAATGGCTA 634

Qy        534 GTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGG 593
              |||||||| |  || | |||||||| || |||||||  |||  || ||||| ||| |||
Db        633 GTGATTCTAAATTGATAAACTTGGCATTGAGAGATTGTGATTTTGTGTTTGATGGATTGG 574

Qy        594 AATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCTT 653
              |||| |||||||||||||||||||||| ||||| ||| |  || ||||| |||||| |||
Db        573 AATCAATTGTGGATGTTGGTGGTGGAAATGGAACCACTGGTAAAATTATTTGTGAGACTT 514

Qy        654 TTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAA 713
              |||||||| |||||| || ||||| ||  | ||| | ||||| || || || || ||||
Db        513 TTCCTAAGTTGAAATGTATTGTGTTTGATAGGCCACAAGTTGTAGAGAACTTATCTGGAA 454

Qy        714 GCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAG 773
              |||| |||||||| | |||||||||||||||||| | ||  || || || ||||||||||
Db        453 GCAATAATTTGACTTATGTTGGTGGGGACATGTTCACATCTATTCCTAATGCTGATGCAG 394

Qy        774 TTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAA 833
              || |||||||||   |||||||||||||| ||| || |||||||  | |||| |  | |
Db        393 TTTTGCTTAAGTATATTTTACATAATTGGACTGATAAGGATTGCACAAGGATACTGAAGA 334

Qy        834 ATTGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTG 893
              | |||||||||||| || ||   ||| | |||| ||| || ||  || | |||||||  |
Db        333 AATGTAAAGAAGCTGTTACAAATGATGGGAAAAAAGGGAAGGTGATTATTATAGATATGG 274

Qy        894 TGATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTAC 953
              | ||||| ||||| |||||||||   |||||||| || | ||||||||||||||||||
Db        273 TCATAAATGAAAAGAAAGATGAGAATCAAGTTACTCAAATTAAGCTCCTTATGGATGTAA 214

Qy        954 ACATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCA 1013
              |||||| ||   | |||||||||||||||||| || ||||| ||||||||||||||||
Db        213 ACATGGCTTG---TCTTAATGGAAAAGAGAGAAATGAGGAAGAATGGAAGAAACTCTTCA 157

Qy       1014 TGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTG 1073
              | ||||| || ||||||  ||| || |||||||| || || |||| ||||||||||||||
Db        156 TAGAAGCTGGTTTCCAAGACTATAAGATATCTCCTTTGACTGGATTTTTGTCTCTTATTG 97

Qy       1074 AGATCTATCCTTGAATACTGACGCTGCAATATTCCATTTAGTAGTTAATTTGCATGTTAT 1133
              |||||||||| | || |||   |      ||    |||  | |   | ||| |||| |||
Db         96 AGATCTATCCATAAACACTTTTGTATTGATCCATCATGCATTCCATTGTTTACATGCTAT 37

Qy       1134 CAATAAA 1140
              || | |
Db         36 AAAGATA 30
```

RESULT 12
BM178815
LOCUS BM178815 546 bp mRNA linear EST 06-DEC-2001
DEFINITION saj59h07.y1 Gm-c1072 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1072-4190 5' similar to TR:O22555 O22555 O-METHYTRANSFERASE. [2] TR:O22308 ;, mRNA sequence.
ACCESSION BM178815
VERSION BM178815.1 GI:17402033
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 546)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 421.

FEATURES Location/Qualifiers
source 1. .546
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1072-4190"
/tissue_type="seedlings induced for symptoms of SDS (Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/clone_lib="Gm-c1072"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library wms constructed from mRNA isolated from 2-3 week old seedlings that were induced for symptoms of SDS (Sudden Death Syndrome) disease by the translocation of culture filtrate of Fusarium solani f. sp. glycines (Plant Cell Report 18:375-380). Cultivar PI 567374 is partially resistant to the disease SDS. Plant tissue (expanded leaves, folded leaves, and new shoots) were collected at 1, 6, 24, and 48 hrs. after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 165 a 101 c 113 g 167 t
ORIGIN

Query Match 35.6%; Score 413; DB 12; Length 546;
Best Local Similarity 85.8%; Pred. No. 2.1e-79;
Matches 471; Conservative 0; Mismatches 75; Indels 3; Gaps 1;

```
Qy        173 ATTACTTTTTCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTG 232
              |||||| || | |||||||||||||||||||| | |||||| |||||  | |||| |||
Db          1 ATTACTCTTCCTGAGTTGGTGTCAATTCTACAAATTCCACCAGCTAAAGTTAGTCAAGTG 60

Qy        233 CAGAGCCTCATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATGAC 292
              || || |||||||| || || | |||||||||||||||||| | ||||||||||||||||
```

```
Db         61 CAAAGTCTCATGCGCTACCTGGAACACAATGGATTCTTTGAAAGAGTAAGAATCCATGAA 120

Qy        293 AACATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTT 352
              ||   ||||||||||||||||||||||||||||||||||| ||||||||||||||| ||||||||
Db        121 AA---AGAAGCATATGCTCTCACTGCTGCTTCAGAGCTACTTGTCAAAAGCTGTGAGCTT 177

Qy        353 AGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAG 412
              |||||||||||||||| |||||| | ||||||| |||| ||    |  | |||   | ||
Db        178 AGTTTAGCTCCAATGATTGAGTTTGTTCTTGACCCAACTTTGTCAAATTCATTCCATCAA 237

Qy        413 TTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCT 472
              ||||| | |||||||| ||||| |||||||||||| |||||||  |||||||||||| || |
Db        238 TTGAAAAAGTGGGTTTATGAGAAAGATCTCACATTATTTGACATCTCCTTAGGATCACAT 297

Qy        473 TTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCT 532
              || |||||||||| | |||||| |||||||  ||||||||||||||||||||||||||||
Db        298 TTGTGGGACTTTCTTAATAAAAACCCTGCACTTAACAAGTCATTCAATGAGGCAATGGCT 357

Qy        533 TGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTG 592
              ||||||||||||||| ||||||||||||| ||||||||||||||||||||| | ||||||
Db        358 AGTGATTCTCAGATGATGAACTTGGCGTTGAGAGATTGCAATTGGGTCTTTCAAGGACTG 417

Qy        593 GAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCT 652
              |||| |||||||||||||||||||||||||||||| |||||| |||||||| ||||| ||
Db        418 GAATTCATTGTGGATGTTGGTGGTGGAACTGGAACCACAGCCAAGATTATATGTGAAGCA 477

Qy        653 TTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGA 712
              |||||||| |||||||||| ||||| || |||||| | ||| | || || |||||||||
Db        478 TTTCCTAACTTGAAATGCATTGTGTTTGACCGTCCACAAGTTATAGAGAACTTGTCAGGA 537

Qy        713 AGCAACAAT 721
              |||||||||
Db        538 AGCAACAAT 546
```

RESULT 13
BQ165418

```
LOCUS       BQ165418                826 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  EST611275 KVKC Medicago truncatula cDNA clone pKVKC-9A4, mRNA
            sequence.
ACCESSION   BQ165418
VERSION     BQ165418.1  GI:20307784
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 826)
  AUTHORS   VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S.,
            Utterback,T., Cheung,F. and Fraser,C.M.
  TITLE     The Medicago truncatula 'kiloclone' set; ESTs selected and
            re-arrayed from various libraries
  JOURNAL   Unpublished
COMMENT     Contact: VandenBosch K
            Department of Plant Biology
            University of Minnesota
            220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
            Tel: 612 624 2755
            Fax: 612 625 1738
            Email: kvandenb@cbs.umn.edu
            TIGR sequence name: MTNAI04TK Alias Clone name:DSIR-19C17 More
            information is available at: www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .826
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pKVKC-9A4"
                     /tissue_type="mixed tissues"
                     /dev_stage="various stages"
                     /lab_host="XLOLR"
                     /clone_lib="KVKC"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                     was directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in XLOLR cells."
BASE COUNT      253 a    127 c    164 g    282 t
ORIGIN
```

Query Match 35.3%; Score 410; DB 13; Length 826;
Best Local Similarity 72.3%; Pred. No. 1e-78;
Matches 581; Conservative 0; Mismatches 205; Indels 18; Gaps 3;

```
Qy          2 TTTGCTATTATGGCTTCTTCATTAAACAATGGCCGTAAAGCAAGTGAGATTTTTCAAGGT 61
              |||||   | |     ||||  | ||  |||  |||||||||||||||| |||||  |||
Db         29 TTTGCAGGTGTATGGGCTTCTTCAATAAATAACCGTAAAGCAAGTGAAATTTTCAAAGCA 88

Qy         62 CAAGCTCTCTTGTACAAACATTTGCTTGGCTTCATAGATTCTAAGTGTCTAAAATGGATG 121
              |||||||| || |||||| |   ||  ||  |||||||||||| | || ||| ||||||
Db         89 CAAGCTCTATTATACAAAAACATGTATGCTTTCATAGATTCCATGTCTCTTAAATGGTCA 148

Qy        122 GTTGAGCTTGACATACCCGACATAATCCACAGCCATAGCCATGGCCAACCCATTACTTTT 181
              |||||  | ||||||||     || ||| ||      |  |||||| |||| |||||| ||
Db        149 GTTGAAATGGACATACCAACTATTATCTAC------AATCATGGCAAACCAATTACTCTT 202

Qy        182 TCAGAGTTGGTGTCAATTCTACAAGTCCCACCAACTAAAACTCGTCAGGTGCAGAGCCTC 241
              ||| | || || |||||||| ||| | ||| |||| |||      | | ||||||  | |||
Db        203 TCAAACTTAGTTTCAATTCTTCAAATTCCATCAACCAAAGTCGATAACGTGCAGCGTCTC 262

Qy        242 ATGCGTTATCTAGCACACAATGGATTCTTTGAGATAGTAAGAATCCATG---------AC 292
              ||||||    || |||||||||||||||||||||||||||||||    |  || |         |
Db        263 ATGCGTCTCCTTGCACACAATGGATTCTTTGAGATAGTAACGAATCAAGAATTGGAAAAT 322

Qy        293 AACATAGAAGCATATGCTCTCACTGCTGCTTCAGAGTTACTTGTCAAAAGCAGTGAGCTT 352
              |   |||||| |||||||||| | | |||||||| | |||| ||| | | ||| |||
Db        323 GAAGAAGAAGCTTATGCTCTCACAGTTACTTCAGAGCTTCTTGTTAAAGGGACTGAACTT 382

Qy        353 AGTTTAGCTCCAATGGTTGAGTATTTTCTTGAACCAAATTGTCAAGGTGCATGGAACCAG 412
              |||||||||||||||||||| | ||||||| || || | | | | |||   | | |
Db        383 TGTTTAGCTCCAATGGTTGAGTGTGTTCTTGATCCTAAATTTAATGTTTCATTTCATAAT 442

Qy        413 TTGAAGAGGTGGGTTCATGAGGAAGATCTCACAGTATTTGAGGTCTCCTTAGGAACACCT 472
              || |||| |||| || |||||||||||||||| | ||||     ||| ||||||      |
Db        443 TTCAAGAAGTGGATTTATGAGGAAGATCTCACCCTCTTTGCTACCTCTTTAGGATGTGAT 502

Qy        473 TTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCT 532
              || ||||| ||| | || ||| | |||| |||||| |   |  |||||||| || ||||||
Db        503 TTATGGGAGTTTCTTAACAAAAATCCTGAATATAATAGATTGTTCAATGATGCTATGGCT 562

Qy        533 TGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTG 592
              |||||||  | ||| | || ||||| || | ||||||||||  || |||||||| ||
Db        563 AGTGATTCCAAAATGATTAATTTGGCATTGAAAGATTGCAATTTTGTGTTTGAGGGGTTG 622

Qy        593 GAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACAGCAAAGATTATCTGTGAGGCT 652
              || || ||||||||||||||||||||| |||||   || |  |||||||| |||||| |
Db        623 GATTCAATTGTGGATGTTGGTGGTGGAAATGGAACAACTGGTAAGATTATTTGTGAGACA 682

Qy        653 TTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGA 712
              | ||||||| ||| |||  | || || || |||||||| ||||| || |||||  |  |||
Db        683 TATCCTAAGTTGAGATGTGTTGTTTTTGATCGTCCAAAGGTTGTAGAGAATTTATGTGGA 742

Qy        713 AGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAAT---GCATCCCCAAGGCTGAT 769
              || |||||||||||||    | | |||  |  | |  ||       | || || ||||||
Db        743 AGTAACAATTTGACATTATGTCGGTGGGGGACATGTTCATATCTTGTTCCTAAAGCTGAT 802

Qy        770 GCAGTTCTGCTTAAGTTGGTTTTA 793
              |||||||||||||||| | |||||
Db        803 GCAGTTCTGCTTAAGTGGATTTTA 826
```

```
RESULT 14
AW733746
LOCUS       AW733746                625 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  sk77f10.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1016-10148 5' similar to TR:O22309 O22309
            7-O-METHYLTRANSFERASE. ;, mRNA sequence.
ACCESSION   AW733746
VERSION     AW733746.1  GI:7639423
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 625)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 411.
FEATURES             Location/Qualifiers
     source          1. .625
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1016-10148"
                     /tissue_type="immature flowers of field grown plants"
                     /lab_host="XL10-Gold"
                     /clone_lib="Gm-c1016"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from immature flowers of field grown plants. The cDNA
                     library was prepared using the Stratagene pBluescript II
                     XR library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into XL10-Gold host cells. This library was
                     constructed by Dr. Randy Shoemaker and Dr. John
                     Erpelding."
BASE COUNT      192 a    101 c    154 g    178 t
ORIGIN

  Query Match             34.9%;  Score 405;  DB 9;  Length 625;
  Best Local Similarity   79.4%;  Pred. No. 1.2e-77;
  Matches  493;  Conservative    0;  Mismatches  125;  Indels    3;  Gaps    1;

Qy        452 GAGGTCTCCTTAGGAACACCTTTCTGGGACTTTATCAATAAAGACCCTGCATATAACAAG 511
              ||  ||||||||||      ||| ||||| ||| | || ||| |||||||||| ||| ||
Db          5 GATATCTCCTTAGGGTGTAGTTTATGGGATTTTCTTAACAAAAACCCTGCATACAACGAG 64

Qy        512 TCATTCAATGAGGCAATGGCTTGTGATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGC 571
              |||||||||||||||||||||| | |||||||| |||    || ||||| || |||||||||
Db         65 TCATTCAATGAGGCAATGGCTAGAGATTCTCAAATGAGTAATTTGGCATTGAGAGATTGC 124

Qy        572 AATTGGGTCTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAATCACA 631
              || | ||| ||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        125 AAGTTGGTGTTTGAGGGACTGGAATCCATTGTGGATGTTGGTGGTGGAACTGGAGCCACT 184

Qy        632 GCAAAGATTATCTGTGAGGCTTTTCCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAAT 691
              || | ||| |||| ||| || |||||| |  |||||||| | ||| | || |||||  |
Db        185 GCCAGGATGATCTCTGAAGCATTTCCTGACTTGAAATGCGTTGTGCTTGACCGTCCTCAC 244

Qy        692 GTTGTGGAAAATTTGTCAGGAAGCAACAATTTGACATTTGTTGGTGGGGACATGTTTAAA 751
              ||| |||| || ||||| | |||||||||||||||||| ||||||||||||||||||| |||
Db        245 GTTCTGGAGAACTTGTCCGAAAGCAACAATTTGACATATGTTGGTGGGGACATGTTCAAA 304

Qy        752 TGCATCCCCAAGGCTGATGCAGTTCTGCTTAAGTTGGTTTTACATAATTGGAATGACAAC 811
              |  || || |||||||||||||||||||||||| | || | ||| |||||| ||| ||
Db        305 TCTATTCCTAAGGCTGATGCAGTTCTGCTTAAGTGGATTCTTCATGATTGGACTGATAAA 364

Qy        812 GATTGCATGAAGATATTAGAAAATTGTAAAGAAGCTAT---TTCAGGTGAAAGCAAAACA 868
              || ||||| |||||||||| |||||||||||||||||    ||||  | |  |||||| |
Db        365 GACTGCATAAAGATATTAGAGAATTGTAAAGAAGCTATTTCTTCAAATAATGGCAAAAGA 424

Qy        869 GGAAAAGTAGTTGTCATAGATACTGTGATAAACGAAAACAAAGATGAGCGCCAAGTTACT 928
              |||||  || | ||||||||||  |||||| | || ||  |||||||||   |||||||
Db        425 CGAAAAATAATCGTCATAGATATGGTGATACAAGAGAAGCAAGATGAGCATAAAGTTACA 484

Qy        929 GAACTAAAGCTCCTTATGGATGTACACATGGCATGTATTATTAATGGAAAAGAGAGAAAA 988
              |||||||||| |||   ||||||  |||||||||| |  |  ||||| |||||||| |
Db        485 GAACTAAAGCTGCTTTGTGATGTAGCCATGGCATGTGTATTATATGGACAAGAGAGAGAT 544

Qy        989 GAAGAAGATTGGAAGAAACTCTTCATGGAAGCAGGGTTCCAAAGCTACAAAATATCTCCC 1048
              || ||||| || |||| ||||| |||||||||||||  | | |  ||||||  |||||| |
Db        545 GATGAAGACTGTAAGAGACTCTACATGGAAGCAGGCGTTCCAGACTACAATATATCTGCA 604

Qy       1049 TTCACAGGATATTTGTCTCTT 1069
                | || || | | ||||||| .
Db        605 CTGACCGGGTCTCTGTCTCTT 625


RESULT 15
BQ165420/c
LOCUS       BQ165420                775 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  EST611277 KVKC Medicago truncatula cDNA clone pKVKC-9A5, mRNA
            sequence.
ACCESSION   BQ165420
VERSION     BQ165420.1  GI:20307788
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S.,
            Utterback,T., Cheung,F. and Fraser,C.M.
  TITLE     The Medicago truncatula 'kiloclone' set; ESTs selected and
            re-arrayed from various libraries
  JOURNAL   Unpublished
COMMENT     Contact: VandenBosch K
            Department of Plant Biology
            University of Minnesota
            220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
            Tel: 612 624 2755
            Fax: 612 625 1738
            Email: kvandenb@cbs.umn.edu
            TIGR sequence name: MTNAI05TV Alias Clone name:DSIR-19L16 More
            information is available at: www.medicago.org
            Seq primer: (gtA AtA CgA CtC ACt AtA ggg C).
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
```

/cultivar="A17"
/db_xref="taxon:3880"
/clone="pKVKC-9A5"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XLOLR"
/clone_lib="KVKC"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

BASE COUNT 272 a 158 c 90 g 255 t
ORIGIN

Query Match 34.9%; Score 405; DB 13; Length 775;
Best Local Similarity 77.6%; Pred. No. 1.2e-77;
Matches 516; Conservative 0; Mismatches 145; Indels 4; Gaps 2;

```
Qy        477 GGGACTTTATCAATAAAGACCCTGCATATAACAAGTCATTCAATGAGGCAATGGCTTGTG 536
              |||| ||| |  ||||| | |||| |||||| |  ||||| ||||| ||||||||| |||
Db        775 GGGATTTTCTTGATAAAAATCCTGAATATAATAGATCATTTAATGATGCAATGGCTAGTG 716

Qy        537 ATTCTCAGATGTTGAACTTGGCGTTTAGAGATTGCAATTGGGTCTTTGAGGGACTGGAAT 596
              |||||  |  || | ||||||||| || ||||||||  |||  || ||||| ||| ||||||
Db        715 ATTCTAAATTGATAAACTTGGCATTGAGAGATTGTGATTTTGTGTTTGATGGATTGGAAT 656

Qy        597 CCATTGTGGATGTTGGTGGTGGAACTGGAATCAC-AGCAAAGATTATCTGTGAGGCTTTT 655
              | ||||||||||||||||||||||||| ||||| |||  |  || ||||| |||||| |||||
Db        655 CAATTGTGGATGTTGGTGGTGGAAATGGAACCACTGGGTAAAATTATTTGTGAGACTTTT 596

Qy        656 CCTAAGCTGAAATGCATGGTGTTGGAACGTCCAAATGTTGTGGAAAATTTGTCAGGAAGC 715
              |||||| ||||||| || ||||| ||  | ||| | ||||| || || || || ||||||
Db        595 CCTAAGTTGAAATGTATTGTGTTTGATAGGCCACAAGTTGTAGAGAACTTATCTGGAAGC 536

Qy        716 AACAATTTGACATTTGTTGGTGGGGACATGTTTAAATGCATCCCCAAGGCTGATGCAGTT 775
              || |||||||| | |||||||||||||||||||| | ||  || || || ||||||||||||
Db        535 AATAATTTGACTTATGTTGGTGGGGACATGTTCACATCTATTCCTAATGCTGATGCAGTT 476

Qy        776 CTGCTTAAGTTGGTTTTACATAATTGGAATGACAACGATTGCATGAAGATATTAGAAAAT 835
              |||||||||    |||||||||||||||| ||| || |||||||  | |||| |  | ||
Db        475 TTGCTTAAGTATATTTTACATAATTGGACTGATAAGGATTGCACAAGGATACTGAAGAAA 416

Qy        836 TGTAAAGAAGCTATTTCAGGTGAAAGCAAAACAGGAAAAGTAGTTGTCATAGATACTGTG 895
              |||||||||||| || ||  |||  | |||| ||| || ||  || | |||||||  ||
Db        415 TGTAAAGAAGCTGTTACAAATGATGGGAAAAAAGGGAAGGTGATTATTATAGATATGGTC 356

Qy        896 ATAAACGAAAACAAAGATGAGCGCCAAGTTACTGAACTAAAGCTCCTTATGGATGTACAC 955
              ||||| ||||| |||||||||   |||||||||| || | ||||||||||||||||| ||
Db        355 ATAAATGAAAAGAAAGATGAGAATCAAGTTACTCAAATTAAGCTCCTTATGGATGTAAAC 296

Qy        956 ATGGCATGTATTATTAATGGAAAAGAGAGAAAAGAAGAAGATTGGAAGAAACTCTTCATG 1015
              ||||| ||   | |||||||||||||||||| || ||||| ||||||||||||||||
Db        295 ATGGCTTG---TCTTAATGGAAAAGAGAGAAATGAGGAAGAATGGAAGAAACTCTTCATA 239

Qy       1016 GAAGCAGGGTTCCAAAGCTACAAAATATCTCCCTTCACAGGATATTTGTCTCTTATTGAG 1075
              ||||| || ||||||  ||| || |||||||| || || |||| ||||||||||||||||
Db        238 GAAGCTGGTTTCCAAGACTATAAGATATCTCCTTTGACTGGATTTTTGTCTCTTATTGAG 179

Qy       1076 ATCTATCCTTGAATACTGACGCTGCAATATTCCATTTAGTAGTTAATTTGCATGTTATCA 1135
              |||||||| | || |||   |     ||    |||  | |   |  ||| |||| ||| |
Db        178 ATCTATCCATAAACACTTTTGTATTGATCCATCATGCATTCCATTGTTTACATGCTATAA 119

Qy       1136 ATAAA 1140
              | | |
Db        118 AGATA 114
```

Search completed: November 10, 2003, 01:51:24
Job time : 2919 secs